

UNIVERZITA KARLOVA

Filozofická fakulta

Ústav bohemistických studií

Bakalářská práce

Vladimír Goroshko

Distribuce a význam slovotvorného sufixu -ák v českých korpusech

Distribution and Meaning of the -ák Suffix in Czech Corpora

Praha 2016

Vedoucí práce: PhDr. Ana Adamovičová

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

V Praze, dne 22. prosince 2016

.....
Vladimír Goroshko

Klíčová slova (česky)

slovotvorba, slovotvorný sufix, sufix -ák v češtině, univerbizace, korpusová lingvistika

Klíčová slova (anglicky)

word building, -ák suffix in the Czech language, univerbisation, corpus linguistics

Poděkování

Rád bych poděkoval vedoucí bakalářské práce PhDr. Aně Adamovičové za odbornou pomoc, cenné rady a připomínky při konzultacích a vypracování bakalářské práce.

Abstrakt (česky)

Práce je zaměřena na teoretické zkoumání vlastností slovotvorného českého sufixu -ák a zjištění statistických zákonitostí jeho distribuce v korpusech českého jazyka. K problému se přistupuje prostřednictvím obecného popisu způsobů tvoření slov v češtině a teoretické klasifikace sufixu -ák podle sémantických a slovotvorných kategorií a typů odvozování. Také se popisují jeho vlastnosti: homonymita, produktivita a citové zabarvení. Praktické zkoumání je zaměřeno na porovnání dat z korpusu psaného jazyka *SYN2010* a korpusu mluveného jazyka *ORAL2013* a vyhodnocení jejich vhodnosti pro výzkum aktuálního jazyka. Průzkum potvrdil, že korpusová data, až na několik málo výjimek popsanych v textu, ve značné míře odpovídají výchozí hypotéze o distribuci tohoto sufixu a že distribuce sufixu -ák vykazuje rozdíly v psaném a mluveném korpusu. V teoretické i v praktické části se zvláštní pozornost věnuje univerbizaci.

Abstract (in English)

The bachelor thesis deals with theoretical exploration of characteristics pertaining to the Czech suffix -ák and determination of statistical correlations relating to its distribution in Czech language corpora. The problem is approached through a general description of word formation in the Czech language and theoretical attribution of the -ák suffix to semantical and morphological categories and derivation paradigms. Particulars of the -ák suffix, such as homonymy, productivity and expressivity, are being described. Practical research is aimed at comparison of data obtained from the *SYN2010* written language corpus and the *ORAL2013* spoken language corpus and assessment of their appropriateness for investigation of the modern language. The research has confirmed that corpus data to a large extent conform with theoretical distribution of the suffix, with a few minor exceptions described in the text, and that such distribution shows certain differences between the written and the spoken corpora. Particular attention to univerbisation is paid both in the theoretical and the practical parts.

OBSAH

1	Úvod	7
2	TEORETICKÁ ČÁST	8
2.1	Způsoby tvoření substantiv v češtině	8
2.1.1	Odvozování (<i>derivace</i>)	9
2.1.2	Mutace, transpozice a modifikace	11
2.1.3	Přípony a jejich kategorizace	11
2.1.4	Homonymita	12
2.1.5	Produktivita	14
2.1.6	Citové zabarvení sufixu <i>-ák</i>	15
2.1.7	Možné důvody expresivity sufixu <i>-ák</i>	16
2.2	Třídění sémantické	17
2.3	Třídění podle slootovorných kategorií (názvy obyvatelské, nositelů vlastností, prostředků apod.)	17
2.3.1	Názvy činitelské	18
2.3.2	Názvy konatelské	19
2.3.3	Názvy prostředků	19
2.3.4	Názvy nositelů vlastností	20
2.3.5	Názvy obyvatelské	21
2.3.6	Názvy příslušenské	22
2.3.7	Augmentativa	22
2.3.8	Jména označující zvířata (<i>přechýlená od feminin</i>)	23
2.4	Třídění podle typu odvozování	23
2.5	Univerbizace	24
3	PRAKTICKÁ ČÁST	27
3.1	Polyfunkční sufix <i>-ák</i> v českých korpusech ČNK	27
3.1.1	Pojem „lemma“	27
3.1.2	Použité nástroje	27
3.2	Pořizování a statistické zpracování výsledků	28
3.2.1	Obecné charakteristiky korpusového dotazu a jeho struktura	29
3.2.2	Dotazy pro vyhledávání v korpusech	29
3.2.3	Frekvenční distribuce	30
3.2.4	Úprava výsledků	30
3.2.5	Přehled nejvíce frekventovaných slov	32
3.2.6	Třídění výsledků	36
3.3	Prezentace výsledků – třídění sémantické (člověk, zvíře, věc, místo, abstraktum)	36
3.4	Prezentace výsledků – třídění podle slootovorných kategorií (názvy obyvatelské, vlastností, prostředků apod.)	38

3.5	Prezentace výsledků – třídění podle typu odvozování (deverbativa, desubstantiva, deadjektiva apod.).....	39
3.6	Prezentace výsledků – statistická distribuce mezi slootovými a sémantickými kategoriemi	40
3.7	Prezentace výsledků – distribuce univerbace v sémantických kategoriích.....	43
3.8	Zajímavý fakt.....	44
4	ZÁVĚR	45
5	SEZNAM POUŽITÉ LITERATURY	47

PŘÍLOHA 1 VÝSLEDEK HLEDÁNÍ V KORPUSU SYN2010

PŘÍLOHA 2 VÝSLEDEK HLEDÁNÍ V KORPUSU ORAL2013

1 Úvod

Sufix -ák je jedním z nejvíce produktivních sufixů v češtině. Nalezneme ho i v jiných slovanských jazycích, nicméně v porovnání s dalšími jazyky jeho rozšířenost v češtině můžeme považovat za velmi vysokou. Vyskytuje se jak ve spisovném, tak i v běžně mluveném jazyce.

V čem spočívá flexibilita tohoto slovotvorného prostředku, schopného zároveň tvořit úplně spisovná slova a vulgarismy? Jaké jsou důvody toho, že pomocí tohoto sufixu snadno vytvoříme výstižný neologismus, vhodný pro daný kontext a bez problémů pochopitelný pro recipienta? Má tento sufix určitá významová omezení?

S ohledem na tyto otázky, cílem dané práce je seznámení s vlastnostmi sufixu -ák a pokus o vysvětlení jeho vysoké frekvence mezi uživateli jazyka na základě dat získaných z korpusů českého jazyka.

Pro účely této práce jsem použil data, která poskytuje projekt Český národní korpus (ČNK). Korpusy ČNK jsou opatřeny nástroji pro vyhledání jednotlivých tvarů v reálných textech a statistické zpracování nalezených výsledků. Projekt nabízí jak korpusy psaného jazyka, tak i mluveného, což je velmi vhodné s ohledem na obecně dominující hovorový charakter sufixu -ák.

V první části práce jsou uvedeny teoretické základy a popsány rozličné kategorie, které mohou být použity pro klasifikaci sufixu -ák. V druhé části jsou tyto poznatky využity pro zjištění distribuce substantiv se sufixem -ák v jednotlivých slovotvorných kategoriích a pro analýzu této distribuce jak v rámci jednotlivých korpusů, tak i mezi různými korpusy.

2 TEORETICKÁ ČÁST

V této části práce se budeme zabývat teoretickými východisky práce, jež jsou podklady pro zkoumání statistických dat, jejich pořízení a zpracování, které je popsáno v druhé, praktické části. Jsou to především způsoby tvoření slov v češtině, specifika sufixu -ák, různé způsoby třídění – sémantické, podle slovotvorných kategorií a typu odvozování a konečně je pozornost věnována procesům univerbizace.

2.1 Způsoby tvoření substantiv v češtině

Nová slova se v jazyce tvoří na základě slov již existujících. Základními a nejrozšířenějšími slovotvornými způsoby jsou odvozování (derivace), skládání (kompozice) a zkracování (abreviace).

Odvozování, při němž se spojuje slovotvorný základ výchozího (fundujícího) slova se slovotvorným formantem, je pro češtinu nejběžnější způsob tvoření nových slov. Pojem odvozování je pro tuto práci klíčový, a tak se mu věnuje větší pozornost v oddílu 2.1.1 *Odvozování (derivace)* níže. Zde uvedeme krátkou charakteristiku dvou dalších způsobů tvoření slov.

Při vzniku slov skládáním se nové slovo tvoří spojením dvou a více slovotvorných základů, například: *velk-* + *-o-* + *-město* → *velkoměsto*, *bíl-* + *-o-* + *-červen-* + *-o-* + *-modrý* → *bíločervenomodrý*, *země-* + *-koule* → *zeměkoule*, *cukr-* + *-o-* + *-vařit* → *cukrovar*, *ok-* + *-a-* + *mžik* → *okamžik*). Výsledek skládání nazýváme kompozitum či složenina.

Rozlišují se kompozice vlastní, při níž výslednou složeninu nelze bez formální změny rozdělit na samostatná motivující slova (*velkoměsto*, *horolezec*, *vodoměr*) a nevlastní, při níž výslednou složeninu lze bez formální obměny rozdělit na výchozí slovní spojení (*okamžik*, *láskyplný*, *chvályhodný*). Zvláštním druhem kompozit jsou příslovečné spřežky, tedy příslovce, vzniklá splynutím předložky s jménem (*nahoru*, *dokonce*, *vtom*).

Zkracováním rozumíme tvoření zkratk (zkratkových slov) ze slov již existujících. Zkratková slova vznikají spojením počátečních hlásek nebo slabik víceslovného názvu, například: *apod.*, *ČR* (Česká republika), *Čedok* (Česká dopravní kancelář). Zkratky dělíme na zkratky čistě grafické, graficko-fonické a fonické.

Zkratky grafické se používají v psaném projevu, zatímco při čtení vyslovujeme jejich plnou formu, a dělí se na iniciálové, jež vznikají z prvních písmen jednotlivých slov víceslovného názvu: *prof.* (profesor), *ing.* (inženýr); kontrakční či stažené, jež jsou utvořené z prvního a posledního písmena: *dr.* (doktor), *pí* (paní); a skeletové, jež jsou složené z charakteristických písmen celého slova: *kpt.* (kapitán), *rtg.* (rentgen).

Zkratky graficko-fonické (také nazývané akronyma) se tvoří z iniciálních hlásek psaných iniciálami. Čteme je buď jako nezkrácené sousloví, nebo po jednotlivých písmenech, která se dají vokalizovat: ČNB (Česká národní banka), FFUK (Filozofická fakulta Univerzity Karlovy), UNESCO (United Nations Educational, Scientific and Cultural Organization). Zkratky fonické mají po stránce zvukové a tvaroslovné charakter slova: *kilo*(gram), *bezva*(dný), *limo*(náda).

2.1.1 Odvozování (*derivate*)

Odvozováním rozumíme způsob tvoření slov pomocí předpon (prefixů), přípon (sufixů) nebo koncovek a jejich kombinací (Hauser, 1980, s. 97). Odvozování je nejčastějším způsobem tvoření slov v češtině a zároveň způsobem, který má největší význam pro cíle této práce, poněvadž právě tímto způsobem jsou tvořena slova vznikající pomocí sufixů.

Pokud jde o substantiva, rozlišuje se odvozování připojením předpony – prefixální (*před-* + *mluva* → *předmluva*), připojením přípony – sufixální (*uči-* + *-tel* → *učitel*), připojením předpony a přípony – prefixálně-sufixální (*na břehu* → *nábřeží*), nahrazením slovotvorné přípony mluvnickou – konverzní (*lovit* + *-0-* → *lov*), připojením předpony a nahrazením slovotvorné přípony mluvnickou – prefixálně-konverzní (*na těle* → *nátělník*, *pa-* + *-hor-* + *-ek* → *pahorek*), odsunutím přípony – desufixální (*dareb-ný* + *-a* → *dareba*), odsunutím přípony a připojením nové přípony – resufixální (*bohatý* → *boháč*) a pouhým odsunutím přípony (v básnickém jazyce) – deprefixální (*útes* → *tes*) (Mitter, [b.r.], s. 14).

Jak již bylo zmíněno, substantiva se sufixem -ák jsou tvořena odvozováním a při další kategorizaci je můžeme zařadit do několika podkategorií: vyskytují se zde typy prefixální (*předškolák*, *nekuřák*, *ultrapravičák*), sufixální (*věšák*, *tučňák*, *spacák*), prefixálně-konverzní (*nástřešák*) a hlavně resufixální (*foťák*, *blondák*, *prvák*).

Při odvozování je důležité rozlišit slova utvořená (fundovaná) a slova neutvořená (prvotní). Utvořené slovo je odvozeno od slova prvotního, čímž s ním vstupuje do základního slovotvorného vztahu zvaného fundace (Hauser, 1980, s. 100).

Slovotvorný základ je část slova, již utvořené slovo má společné se slovem základovým. Řadu příbuzných slov se stejným základem nazýváme slovní čeledí. Například, ve slovní čeledi „les, lesní, lesnatý, zálesí“ je slovotvorným základem prvek „les“. Slovotvorný základ se může krýt s kořenem základového slova, ale v mnoha případech dochází před připojením té či jiné přípony k dalším změnám, například odstranění koncovky (*škola* – *školní*), anebo změně hlásky (*slibit* – *slib*, *stůl* – *stolek*, *pes* – *psí*) (Hauser, 1980, s. 101). Hláskové změny mohou mít rozsáhlý charakter. Tak, pokud jde o samohlásky, může dojít k jejich krácení (*ráno* → *ranní*), dloužení (*hrad* → *hrádek*), střídání (*schod* → *schůdek*), výměně dvojhlásky za samohlásku (*brousit* → *brusný*), vzniku nebo zániku (*jádro* → *jadérko*, *pes* → *psí*). Pokud jde o souhlásky, může dojít k jejich měkčení (*jih* → *jižní*, *ruka* → *ručka*) anebo tvrdnutí (*běžet* → *běh*). Takové hláskové realizace jednoho morfému nazýváme alomorfy.

Prvek, pomocí něhož se odvozené slovo tvoří od prvotního slova, se nazývá formant nebo afix (Hauser, 1980, s. 104). Funkci formantu může plnit předpona, přípona, koncovka nebo jejich kombinace.

Při odvozování pomocí předpony (prefixace) se tvaroslovný charakter slova nemění a slovní druh odvozeného slova zůstává stejný (*táhnout* → *natáhnout*, *poručík* → *podporučík*, *starý* → *prastarý*). Odvozování pomocí přípony (sufixace) může probíhat jak beze změny slovního druhu (*chlap* → *chlapík*, *divný* → *divnější*), tak i s takovou změnou (*učit* → *učitel*, *kniha* → *knižní*). Jak je možné vidět z posledního příkladu, přípona v sobě obvykle zahrnuje i koncovku. Koncovka však může sloužit i jako samostatný formant (konverze) (*kámen* → *kamení*, *chválit* → *chvála*, *lovit* → *lov*). Existuje také smíšený způsob odvozování, při kterém se používá formantů různého druhu (*ob* + *oči* + *i* → *obočí*, *ná* + *čelo* + *ník* → *náčelník*).

Méně častými typy afixů jsou postfix, infix a interfix. Postfix je morfém, jenž se připojuje úplně na konec slova za koncovku. V češtině ho najdeme jen u některých neurčitých zájmen a příslovců (*kdosi*, *kdekoli*, *kdepak*). Infix je morfém, jež se vkládá do slovního základu a nejčastěji slouží pro intenzifikaci míry vlastnosti vyjadřované adjektivem či adverbiem (*mal-*

in-ko). Interfixem nazýváme spojovací morfém mezi dvěma slovními základy složeného slova: *sever-o-západ*, *smysl-u-plný*.

2.1.2 *Mutace, transpozice a modifikace*

Proces tvorby substantiv může mít mutační, transpoziční anebo modifikační charakter.

Mutací nazýváme slovotvorný proces, při němž vzniká nové slovo, které označuje jinou skutečnost než slovo základové (David, 1999, s. 16). Při mutaci se význam slova fundovaného liší od významu slova prvotního a může docházet ke změně slovního druhu, například: *milovat* – *milovník*, *červený* – *zčervenat*, *pošta* – *pošťák*.

Při transpozici je význam prvotního a fundovaného slova totožný a mění se jen slovnědruhovká kategorie. V této kategorii se substantiva se sufixem -ák nevyskytují, příklady s jinými substantivy jsou: *mladý* – *mladost*, *cizí* – *cizost*, *lovit* – *lov*.¹

Při modifikaci je význam slova fundovaného pouze modifikován ve srovnání s významem slova prvotního. Patří sem deminutiva (*domek*, *chlapeček*, *maminka*), augmentativa (*psisko*, *chlapák*, *babizna*), komparace (*velký* – *větší* – *největší*, *dobrý* – *lepší* – *nejlepší*) a přechylování (*žákyně*, *úřednice*, *myšák*). V případě sufixu -ák se v této kategorii setkáme především s augmentativy (*chlap* – *chlapák*, *auto* – *auták*, *šutr* – *šutrák*) a jmény mužského rodu, která označují zvířata a jsou přechýlená od feminin (*myš* – *myšák*, *husa* – *husák*, *opice* – *opičák*).

2.1.3 *Přípony a jejich kategorizace*

Přípona (sufix) je slovotvorný prostředek, který se připojuje za slovotvorný základ. Tento morfém nemá v jazyce samostatnou existenci a vždy je vázán na slovo (Hauser, 1980, s. 104). Přípony rozdělujeme na kmenotvorné, slovotvorné a tvarotvorné. Kmenotvorné přípony slouží pro spojení kmenotvorného morfému s tvarotvornou příponou a vyskytují se u sloves (*nes-e-0*, *tisk-ne-0*, *kry-je-0*, *pros-í-0*, *děl-á-0*) a substantiv typu „kuře“ (*kuř-0-e*, *kuř-et-e*,

¹ Patří sem také konvertovaná slova, která při přechodu z jednoho slovního druhu do druhého zachovávají svoji formu: *nemocný*, *raněný*, *konečná*.

kuř-et-i, kuř-at-a, kuř-at-0). Tvarotvorné přípony slouží k vyjádření gramatických kategorií a dělí se na nefinální (při-nes-0-l-i, u-děl-á-n-0, za-č-n-i-0) a finální (koncovky) (nes-0-u, děl-a-l-a, muž-í, jar-n-í).

Sufix -ák patří do kategorie přípon slovotvorných, které slouží k modifikaci lexikálního významu celého slova nebo slovního tvaru. Slovtvorná přípona často slouží nejen k utváření kmenu odvozeného slova, ale také formuje slovní druh nově tvořeného slova. Podle příslušnosti odvozeného slova k určitému slovnímu druhu se dělí slovotvorné přípony na substantivní (čten-áři, kuř-ák, loď-ka), adjektivní (plat-ný, viditel-ný, psa-cí), příslovečné (úmysln-ě, rychl-ě, snadn-o) a slovesné (sedě-t, plan-uje, chodi-l) (Dokulil, 1962, s. 150). Sufix -ák začleňujeme do přípon substantivních. Podstatná jména tvořená pomocí sufixu -ák mají morfologickou charakteristiku substantiv mužského rodu životného nebo neživotného a zařazují se do vzoru „pán“ nebo „hrad“.

Přípony kategorizujeme podle formální a významové stránky. Formálně je zkoumaný sufix -ák tvořen dvěma hláskami a má pevnou podobu, tj. netvoří alomorfní varianty, což znamená, že se jeho tvar nikdy nemění, respektive mění se jen v rámci deklinace slov mužského rodu životného – v nominativu (-áci) a lokálu plurálu (-ácích). Po významové stránce se však sufix -ák vyznačuje poměrně širokým rozsahem významů, můžeme ho označit za vysoce polyfunkční sufix mající jednoznačně dominantní roli v procesu univerbizace. Jak zjistíme dále při různých typech třídění v rámci dané práce, slouží ke tvoření více významových kategorií (viz oddíly 2.3 *Třídění sémantické*, 2.4 *Třídění podle slovotvorných kategorií*, 2.5 *Třídění podle typu odvozování* a 2.6 *Univerbizace* níže).

2.1.4 *Homonymita*

Typickým rysem přípon obecně a jmenovitě sufixu -ák je poměrně vysoká homonymita. Homonymitou rozumíme stejné znění slov nebo slovních formantů různého významu. Homonymita je spojená se sémaziologií, tj. přístupem ke zkoumání významu slov, při kterém postupujeme od formy k významu.²

² Tento postup je opačný postupu onomaziologickému, při kterém se naopak postupuje od významu ke formě.

Homonymita se prolíná mnohými vrstvami jazyka a setkáváme se s homonymitou lexikální, slovotvornou, slovnědruhovou, morfologickou a syntaktickou. Lexikální homonymita se projevuje na úrovni slov, nap. *kolej* (ubytovací zařízení a jízdní cesta kolejových vozidel), *role* (úloha a svitek), *pero* (pružina a propisovací tužka). Homonymitou slovnědruhovou nazýváme výsledek konverze, tj. přechodu slova do jiného slovního druhu beze změny formy, např. *večer* (substantivum a adverbium), *vrchní* (adjektivum a substantivum). Homonymita morfologická se vyznačuje totožností formy napříč tvarového systému jednoho nebo více slov a najdeme ji u slov jako *je* (3. osoba přítomného času verba „být“ a 4. pád zájmena „oni“), *bratrovi* (3. pád singuláru jména „bratr“ a 1. pád plurálu přivlastňovacího adjektiva „bratrův“ u mužského rodu životného), *jeho* (2. a 4. pád osobního zájmena „on“ a přivlastňovací zájmeno) apod. Homonymita syntaktická je zvláštním případem homonymity a zařazujeme do ní věty, jejichž smysl se může lišit v závislosti na kontextu.

Homonymita slovotvorná nás zajímá nejvíc a v tomto případě jde právě o homonymii morfémů (kořenů a afixů – předpon a přípon). Například, verbum *rozmluvit* tvořené předponou roz- může mít významy „nechat hovořit“ a také „přesvědčit někoho, aby něco nedělal“. Toto poukazuje na homonymitu předpony roz-, která nese více významů. Ve zmíněném případě jsou tyto významy odlišeny jinou valencí verba: v prvním případě *někdo se rozmluvil*, v druhém *rozmluvíme někomu něco*. Pokud jde o substantivní přípony, jmenovitě o sufix -ák, valenční doplnění slova homonymii neodstraní, zjednotnění vzniká až v kontextu: *Dám si multák* vs. *Jdu do multáku*. Příčiny homonymity přípon vysvětluje např. P. Hauser (1980, s. 105):

Homonymita přípon je značná. Je vysvětlitelná omezeným počtem přípon a značným počtem slovotvorných významů. Inventář přípon je v jazyce relativně stálý. Ve vývoji dochází ke změnám ve využívání přípon, v jejich produktivitě, ale málokdy vznikají přípony zcela nové.

Co se týče sufixu -ák, jak je popsáno níže, tento sufix slouží ke tvoření širokého spektra kategorií sémantických (*člověk, zvíře, věc, místo a abstraktum*) a slovotvorných (názvy činitelské, konatelské, názvy prostředků, nositelů vlastností, názvy obyvatelské, příslušenské, augmentativa a jména zvířat přechýlená od feminin). Není toto dělení však definitivní a při další podrobnější kategorizaci bychom mohli třeba jména označující prostředky zařadit do podskupin skutečných nástrojů (*naběrák*), druhů tanců (*dupák*), části zvířecího těla (*běhák*),

v kategorii denumerálií by se daly vyčlenit jména označující pořadí (*prvák, druhák*) a jména označující srnčí zvěř podle paroží (*desaterák, dvanáctérák*).

2.1.5 Produktivita

Přípona -ák patří do produktivních slovotvorných prvků. Za produktivní označujeme morfémy, pomocí kterých se tvoří slova v současném jazyce (Hauser, 1980, s. 111). Kvantita produktivity je sice těžko měřitelná, nicméně je třeba říci, že stupeň produktivity je vyšší u slovotvorných prostředků, které mají větší rozmanitost po stránce sémantické, jinými slovy produktivita stoupá s počtem rozmanitých významů, které vznikají u základů formovaných pomocí toho či jiného sufixu (Dokulil, 1962, s. 85).

O produktivitě sufixu -ák zejména vypovídá ten fakt, že často při běžné komunikaci v češtině narážíme na specifické lexikum, používané jednotlivými mluvčími nebo v profesních kruzích, avšak takovéto lexikum ve slovníku nebo korpusu nenalezneme. Z vlastní zkušenosti mohu uvést například taková slova jako *kraták* (krátké spojení či zkrat), *dvouřadák* (kabát s dvěma řadami knoflíků), *basketák* (basketbalista), *běžňák* (běžný účet) nebo *festák* (festival). Jsou to poměrně běžně používaná slova, která se však nevyskytují v jazykových korpusech *SYN2010* (Křen et al., 2010) a *ORAL2013* (Benešová et al., 2013), jež jsou zdrojem dat pro praktickou část této práce. Příčinou této absence mohou být jak produktivita sufixu -ák, tak i jeho rozšířené použití pro aktuální tvoření nových slov, která se do korpusů hned nedostanou. Budování korpusů je dlouhodobý proces, při němž se snaha o vyváženost zdrojů podílí na následné míře reprezentativnosti daného korpusu. I při sebelépe nastavených kritériích se nově vznikající slova nemusí nutně do korpusu dostat, což souvisí jednak s jejich frekvencí užití, jednak s idiolektem toho kterého mluvčího, obzvlášť v případě mluvených korpusů.

V tomto ohledu můžeme položit oprávněnou otázku, zda a do jaké míry moderní korpusy zachycují současný jazyk. Odpověď na tuto otázku však přesahuje rámec dané práce. V souvislosti s tímto problémem si korpusoví lingvisté uvědomují nutnost studia slov, která se v korpusech vyskytla jen jednou, tzv. *hapax legomenon*, protože se právě mezi nimi mohou vyskytnout nová slova v češtině.

V případě sufixu -ak nám korpusy poskytují ještě jeden velmi cenný zdroj informací: studium frekvence jednotlivých stylových dublet v kontextu a možnost určení frekventovanějšího tvaru (*Pražan – Pražák, Vídeňan – Vídeňák* apod.).

2.1.6 Citové zabarvení sufixu -ák

Přidáním sufixu -ák výsledné slovo získává ve většině případů určitý citový příznak. Jde především o zabarvení expresivní až pejorativní. Na rozdíl například od sufixu -ka, slova tvořená pomocí sufixu -ák nesou hrubší maskulinní kvalitu. Vezměme například slova *magnetka* (magnetická střelka) a *magneták* (magnetofon). První vyjadřuje náš pozitivní vztah k danému předmětu a zachycuje jeho malý rozměr, zatímco druhé zní technicky odcizeně a má jistý expresivní odstín.

Jen málo podstatných jmen na -ák nemá žádné citové zabarvení či mluvený ráz (*čmelák, divák, dudák, kuřák, voják, zpěvák, Polák, Slovák* apod.). Také by se dalo říct, že je dnes tento podtyp zcela neproduktivní. Některá substantiva dnes již vnímáme jako neutrální, i když výrazně mluvená slova (*obyvák, řidičák, magneták, foťák*). Mnohem více neutrální jsou však podobně utvořená feminina na -ka (např. *bakalářka, diplomka, cestovka, občanka*). Většina substantiv na -ák vykazuje silnější nebo slabší pejorativní zabarvení, podepřené často již zhrubělým zabarvením základového slova (Dokulil, 1967, s. 72). Například: *drsňák, krvák, hnusák*. Pejorativnost je také výraznější například u slova *Rusák* než *Pražák*, které má spíše pozitivní konotaci. Alternuje-li příponě -ák spisovná varianta -an (*Pražan – Pražák*), nově utvořené slovo nemá pejorativní nádech jako v případě, když sufix -ák přímo přidáme k základu slova (*Rus – Rusák, Žid – Židák*).

Při použití sufixu -ák také můžeme zaznamenat zvýšení míry kvality původního slova. Jako příklad může posloužit slovo *dobrák*, které zdůrazňuje kvalitu „dobroty“, a tak slovo *dobrák* nese význam větší dobroty než sousloví „dobrý člověk“.

Expresivita sufixu -ák se nejvíc projevuje u jmen činitelských (*piják, lajdák*). Expresivní odstín se nicméně rovněž uplatňuje u jmen konatelských (*hrobák, sladák*) a také v menší míře u pojmenování prostředků (*bzučák* (bzučící elektrický zámek), *mokvák* (zmokvalá kůže), (Dokulil, 1967, s. 225).

V mnoha případech také varianty tvořené sufixem -ák, především odvozené od adjektiv a univerbizované, zní vtipně a slouží k pobavení svým výstižným popisem toho či jiného předmětu nebo skutečnosti. Například, slovem *žluták* by se dala označit jakákoliv věc žluté barvy či člověk oblečený do žlutých šatů, a tak v řeči toto slovo často nahrazuje víceslovná pojmenování jako „žlutý autobus“ nebo „chlap ve žlutém“, případně „žlutý kůň“ nebo „žlutý pes“ apod.³

Zvláštním jevem podporujícím expresivitu jsou modifikace kmenové souhlásky, které se zejména projevují jako změkčení (*cesták*, *šprták*, *Židák*) (Dokulil, 1967, s. 73). Slova daného typu prakticky vždy mají znevažující charakter. Občas se ale při alteraci kmenové souhlásky mění i význam celého slova (*světák* (kavalír) – *svěťák* (světový rekord); *smeták* (zametací prostředek) – *smet'ák* (smetiště).

2.1.7 Možné důvody expresivity sufixu -ák

Pro slova tvořená pomocí sufixu -ák je typické, že nemají odvozený význam, a to ani přenesený, ani metaforický. Toto je bezesporu velmi zajímavý rys, který v podstatě zachycuje to, jak je tento slovotvorný prostředek vnímán uživateli jazyka. Ve většině případů totiž jde o konkrétní vlastnost (*držák* je především to, co drží; *koupák* je místo, jehož hlavní funkcí je koupání apod.). Tvořením univerba s příponou -ák je tato vlastnost často ještě víc zdůrazňovaná a nabývá na intenzitě. Tento jev se obzvlášť projevuje u některých slov, odvozených od adjektiv (jako např. už zmíněný *dobrák*, či obdobně např. *hlupák*, který je hloupější než pouze *hloupý*, anebo *doják* bývá dojemnější než pouze *dojemný*).

Myslím si, že tato možnost snadného zdůraznění určité praktické stránky věci či nápadného znaku je právě tím, co tak imponuje uživateli češtiny a je zároveň snadným prostředkem pro tvoření dalších a dalších nových slov pomocí přípony -ák.

Jak již bylo zmíněno, v mluveném projevu se sufixace přidáním sufixu -ák (a podobně také sufixů -ka, -ovka a -átko) často používá za účelem pobavení. Myslím si, že zaručený úsměv

³ Zároveň, je-li slovo psáno velkým počátečním písmenem, se může jednat o přezdívku (jako například u osmi výskytů slova *Žlutá* v Pražském mluveném korpusu – jak jsem se doslechl, jde o variantu přezdívky *Žluták*).

mezi účastníky konverzace vyvolává právě tato možnost snadno a pohotově vystihnout určitou vlastnost krátkým jednoslovným pojmenováním, často utvořeným *ad hoc*. Za svou oblibu tato slova vděčí právě této hravosti a časté expresivitě, ovšem také výrazně ekonomickému charakteru jednoslovných pojmenování v případě univerbizace oproti těžkopádnosti pojmenování víceslovných.

2.2 Třídění sémantické

Při sémantické klasifikaci se slova rozdělují z hlediska jejich věcného významu. S ohledem na významovou škálu sufixu -ák, pro sémantické třídění v rámci této práce byly použity kategorie *člověk*, *zvíře*, *věc*, *místo* a *abstraktum*. Tyto kategorie zahrnují celé spektrum možných významů substantiv se sufixem -ák.

2.3 Třídění podle slovotvorných kategorií (názvy obyvatelské, nositelů vlastností, prostředků apod.)

Slovotvorný typ zahrnuje slova tvořená stejným způsobem za stejných podmínek a se stejným obecným významem (Hauser, 1980, s. 108). Z více slovotvorných typů vznikají slovotvorné kategorie, které zahrnují slovotvorné typy s podobným významem.

Jednotné formanty buď odpovídají jedné slovotvorné kategorii, nebo jsou polyfunkční a mohou být zařazeny do několika tříd. Sufix -ák nese více významů, je funkčně vysoce zatížený a může být přiřazen k následujícím třídám v rámci pojmenovací kategorie mutace: názvy *činitelské*, *konatelské*, názvy *prostředků*, *nositelů vlastností*, názvy *obyvatelské*, *příslušenské*, *augmentativa* a *jména zvířat přechýlená od feminin*. Sufix -ak se nevyskytuje pouze u třídy výsledky děje (*nomina resultativa*) a názvů míst (*nomina loci*).⁴

⁴ U názvů míst se původně sice nevyskytuje, u univerb je však jeho výskyt jednoznačně mohutný: srov. např. *Václavák*, *Staromák* apod.

2.3.1 Názvy činitelské

Názvy činitelské pojmenovávají osoby a zřídka zvířata, rostliny či jevy podle jejich charakteristické činnosti a tvoří se ze slovesných základů (například: *kuřák* kouří, *žebrák* žebře) (David, 1999, s. 17). Jsou to deverbativa.

Obecně jsou názvy činitelské nejčastěji tvořeny následujícími příponami:

- tel (učitel, ředitel, stavitel, velitel, vychovatel)
- č (topič, hlídač, hráč, posluchač, obaleč)
- ce (obránce, vládce, rádce, soudce, obhájce)
- ec (letec, věštec, chodec, běžec, horolezec)
- čí (mluvčí, rozhodčí, průvodčí, výberčí, výpravčí)
- ník (dělník, milovník, loupežník, bojovník, spolupracovník)
- ák** (divák, tulák, žebrák, zpěvák, liják, piják)
- ař (kuchař, pekař, písař, těžař, metař)
- ář (čtenář, kovář, lhář, opravář, řezbář)
- al/-il (čumil, hýřil, kecal, kutil, střádal, tlachal, žvanil)
- an (píjan, trhan, hafan, rafan)
- oun (pěstoun, šplhoun, mlsoun)
- na (drbna, klepna, běhna)
- dlo (fintidlo, trapidlo, trdlo)

Co se týče substantiv činitelského typu se sufixem -ák, tvoří se nejčastěji na základě infinitivního kmene sloves bez kmenové přípony (*zpěvák*, *řezák*, *horák*). Někdy se tvoří i na základě kmene přítomného (*pasák*, *zabiják*, *ležák*) (David, 1999, s. 19).

Kromě některých výjimek, které nemají citový příznak (*divák*, *zpěvák*, *kuřák*), má většina činitelských názvů tvořených příponou -ák pejorativní zabarvení (*ulejvák*, *courák*, *vandrák*) (Hauser, 1980, s. 117).

V některých případech při vzniku názvů činitelských dochází ke krácení samohlásky v základu slova. Tento jev je však jen příležitostný, není dnes produktivní a vyskytuje se jen u starších slov, například: *kuřák*, *tulák*, *zpěvák* (Dokulil, 1967, s. 73).

2.3.2 *Názvy konatelské*

Kategorie názvů konatelských je velmi blízká kategorii názvů činitelských. Na rozdíl od názvů činitelských se substantiva této kategorie tvoří od jiných substantiv, jsou to totiž desubstantiva.

Obecně jsou názvy konatelské nejčastěji tvořeny těmito příponami:

- ař (bruslař, chatař, lékař, lyžař, vorař, lékař, sochař, vinař)
- ář (rybář, hvězdář, brankář, farář, novinář, truhlář, sukničkář)
- ník (lesník, horník, obchodník, kouzelník, úředník)
- ák** (autobusák, ovčák, pytlák, saniták, plamenák, křížák, světák)
- íř (nožíř, hrnčíř, mečíř, uhlíř, zbrojíř)

Uvedme další příklady z řady frekventovaných substantiv se sufixem -ák: *školák, ovčák, květák, voják, mrazák, parťák*.

2.3.3 *Názvy prostředků*

Tato třída je poměrně široká a tvoří se na základě sloves, tj. má deverbativní charakter. Patří do ní prakticky všechny názvy předmětů (ne osob) se sufixem -ák odvozené od sloves.

Obecně jsou názvy prostředků nejčastěji tvořeny následujícími příponami:

- č (nosič, ohřívač, odpařovač, vysoušeč)
- čka (vrtačka, čistička, prohlížečka, přehazovačka, pračka)
- ka (pojistka, výhybka)
- dlo (letadlo, okysličovadlo, lepidlo, rozpouštědlo, mýdlo)
- tko (kukátko, razítko, ukazovátko, struhátko, pravítko)
- ník (narázník, cedník, dopravník, výměník)
- ák** (bodák, hořák, leták, padák, sušák, držák, věšák, šroubovák, otvírák, zobák)
- ivo (krmivo, mazivo)
- ík (budík, hoblík, špendlík)

Jména prostředků na -ák často mají významové ekvivalenty tvořené jiným sufixem (-č, -čka, -dlo, -ka apod.). V těchto případech mají názvy na -ák odstín hovorový až slangový,

například: *utahovák* (*utahovač*), *naběrák* (*naběračka*), *ohřívák* (*ohříváč*) (Dokulil, 1967, s. 225).

2.3.4 *Názvy nositelů vlastností*

Názvy nositelů vlastnosti jsou odvozeny především z adjektiv (*chytrák*), sloves (*uprchlík* – mezistupněm je zde deverbativum *uprchlý*), číslovek (*třetina*), výjimečně substantiv (*slovník*).

Pokud jde o sufix -ák, sem patří většinou označení lidí podle jejich vnějšího znaku nebo charakteristického rysu (*pravák*, *bouřlivák*, *bělák*), označovaná vlastnost se však může vztahovat i na zvíře (*ušák*, *hnědák*, *šesterák*) nebo věc (*širák*, *kožeňák*, *zlaták*) apod. (David, 1999, s. 26). Zpravidla se tato substantiva, pokud jsou tvořená od adjektiv sufixem -ák, vyznačují expresivním nádechem (*hlupák*, *chudák*, *sprosták*) (Hauser, 1980, s. 121). Jak již bylo zmíněno, při tvoření těchto slov dochází ke zvýšení míry odpovídající charakteristické vlastnosti jejího nositele.

Obecně jsou názvy nositelů vlastností nejčastěji tvořeny níže uvedenými příponami:

- ec (stařec, zbabělec, nosorožec, zaměstnanec, poctivec)
- ík (mladík, odborník, loupežník, útočník, rychlík)
- ák (cizák, pravák, levák, pravičák, dobrák, hlupák, chudák, chytrák, paličák)
- ek (mládek, stárek, předek, levoboček, párek)
- áč (boháč, vousáč, roháč, zavináč)
- och (běloch, černoch, lenoch, divoch)
- ouš (bělouš, starouš, chytrouš)
- ýš (slepýš, hroznýš, chřestýš, měkkýš)
- as (chud'as, klid'as, prud'as)
- án (dlouhán, velikán)
- ice (hořčice, rovnice, borovice, slunečnice)
- ka (cukrovka, lanovka, ukolébavka, tlačenka)
- ina (krytina, slanina, uzenina, zavařenina)

M. Dokulil zahrnuje do kategorie nositelů vlastností také univerbizační tvoření citově zabarvených výrazů z dvojčlenných přívlastkových sousloví (*živnosták* ze „živnostenský

list“, *děcák* ze „dětský domov“, *smažák* ze „smažený sýr“ apod.) (Dokulil, 1967, s. 224). Jevu univerbizace se v rámci dané práce však budeme dál věnovat a budeme ho zkoumat zvlášť jako jeden z typu odvozování (viz oddíl 2.6 *Univerbizace* níže).

2.3.5 *Názvy obyvatelské*

Třída obyvatelských názvů zahrnuje názvy osob podle místa, kde bydlí nebo odkud pochází. Ve spisovném jazyce je v této oblasti sufix -ák málo produktivní. Z obecních názvů jsou tvořeny například *skalák*, *horák*, *měšťák* (Hauser, 1980, s. 123). Jsou to desubstantiva a jejich základem jsou buď apelativa (*hora* – *horák*) nebo propria (*Praha* – *Pražák*). K názvu místa nemusí vždy existovat i název jeho obyvatele, jeho existence je podmíněna frekvencí názvu místa a druhem sufixu tohoto názvu (např. *Karlovy Vary*, *Haag*).

Obecně jsou názvy obyvatelské nejčastěji tvořeny následujícími příponami:

- an (Pražan, měšťan, občan, vesničan, měšťan, Maročan)
- ec (cizinec, Portugalec, Filipínek, Kubánek, Korejec)
- ák (lesák, měšťák, Hanák, Eskymák, Dejvíčák, Žižkovák)
- ník (poustevník)
- 0 (Čech, Ind, Rus, Nor, Srb) – desufixace

Zajímavá situace se vyskytuje v oblasti vlastních jmen, jež je možné rozdělit do čtyř typů. U jedněch sufix -ák tvoří zcela spisovné tvary, jako *Eskymák*, *Slovák*, *Polák*. U druhých s jeho pomocí vznikají alternativní hovorové varianty, které v případě posledního příkladu mohou mít i pejorativnější zabarvení: *Pražan* – *Pražák*, *Moravan* – *Moravák*, *Rus* – *Rusák*. Do třetího typu můžeme zařadit vlastní jména, u nichž hovorové varianty stále více pronikají do jazyka spisovného, kde vytlačují příponu -an: *Brňák*, *Videňák*. U čtvrtého typu *měšťan* – *měšťák* dochází k významovému posunu slova. Zatímco *měšťan* je obyvatelem města, *měšťákem* označujeme člověka, který se jen vnějškovými znaky snaží napodobit obyvatele měst. Tato charakteristika je ještě výraznější u kompozita *maloměšťák*. V obou případech jde o přechod mezi třídou obyvatelskou a třídou příslušenskou s přesahem k augmentativům.

2.3.6 *Názvy příslušenské*

Slova patřící do této kategorie označují příslušníky určité zaměstnanecké nebo společenské vrstvy, organizace, politické strany apod. Substantiva tvořená pomocí přípony -ák většinou mají mluvený ráz (*zbrojovák, škodovák, ajťák, levičák, pravičák*). Jsou to desubstantiva mužského rodu.

Obecně jsou názvy příslušenské nejčastěji tvořeny následujícími příponami:

- ník (družstevník, vrstevník, kališník)
- ec (Přemyslovec, ruchovec, lumírovec)
- ář (odborář, odbojář, legionář, revolucionář)
- ák (junák, zbrojovák, škodovák, měšťák, voják)

2.3.7 *Augmentativa*

Předposlední třída, v níž se vyskytují substantiva utvořená příponou -ák jsou augmentativa patřící do pojmenovací kategorie modifikační.

Obecně augmentativa, anebo zveličená jména, jsou nejčastěji tvořena těmito příponami:

- ák (auťák, chlapák, študák, měšťák)
- isko (čertisko, chlapisko, psisko, klučisko, křovisko)
- an (nosan, zoban)
- as (lotras)
- izna (babizna)

Augmentativa nejsou frekventovaným prostředkem českého jazyka. Sufix -ák se nicméně podílí na jejich tvorbě, pravděpodobně kvůli svému pejorativnímu zabarvení, které se shoduje s obecně pejorativním charakterem augmentativ. Dle typu odvozování jsou to desubstantiva.

2.3.8 Jména označující zvířata (přechýlená od feminin)

Poslední třídou, v níž se vyskytuje sufix -ák jsou přechýlené názvy mláďat (*nomina mota*) patřící též do pojmenovací kategorie modifikační.

Obecně jsou přechýlená jména odvozena od maskulin a jsou tvořena příponami:

- ka (sousedka)
- yně (mistryně, ministryně)
- kyně (rádkyně, soudkyně)
- ice (holubice)

Přechýlená jména se sufixem -ák se vyznačují opačným přechylovacím postupem, totiž odvozením maskulina na základě feminina. Sufix -ák je nejrozšířenějším prostředkem pro tvoření podstatných jmen tohoto typu (*lišák, myšák, opičák, žabák*) (Dokulil, 1967, s. 551).

2.4 Třídění podle typu odvozování

Nově tvořená slova můžeme také třídit na základě slovnědruhové příslušnosti základového slova. Tím pádem, jak už bylo naznačeno výše, rozlišujeme odvozeniny od sloves – deverbativa a od jmen – denominativa, které dále lze třídit na odvozeniny od jmen podstatných – desubstantiva a od jmen přídavných – deadjektiva (Hauser, 1980, s. 108). Příklady deverbativ se sufixem -ák: držet – *držák*, šroubovat – *šroubovák*; příklady desubstantiv: *měšťák, obchodák*; příklady deadjektiv: *chudák, chytrák*. Další samostatnou kategorií odvozenin jsou odvozeniny od číslovek, do které řadíme následující substantiva se sufixem -ák: *prvák, druhák, třeták* apod. Tato kategorie sice není uvedena v současných mluvnicích, je ale poměrně frekventovaná a tudíž nezbytná pro třídění reálně se vyskytujících současných korpusových dat v rámci této práce.

Dalo by se říci, že v podstatě každá slovotvorná třída často využívá jako zdroj základová slova patřící do jen jednoho slovního druhu a je vázána na jeden určitý typ odvozování. Jména činitelská jsou vždy deverbativa, zatímco jména konatelská, obyvatelská, příslušenská, augmentativa a jména zvířat přechýlená od feminin jsou vždy desubstantiva. Jména prostředků jsou většinou deverbativa (výjimečně jsou odvozena i od jmenných základů – např. *číslice*), jména nositelů vlastností jsou většinou deadjektiva. Z dalších dvou

tříd, u kterých se přípona -ák nevyskytuje, jména výsledků děje jsou čistá deverbativa, zatímco jména místní mohou být desubstantiva, deverbativa i deadjektiva.

2.5 Univerbizace

Univerbizací rozumíme tvoření substantiv, jež nahrazují pojmenování víceslovné. Nejčastěji se tento jev vyskytuje v mluveném jazyce a slangu, kde jednoslovné názvy vznikají snahou o stručnost vyjádření a jsou projevem jazykové ekonomie.

V kontextu dané práce nás univerbizace především zajímá proto, že mezi typické produktivní odvozovací přípony pro univerbizaci, kterými jsou podle P. Huasera přípony -ář, -ák, -ka/-ovka a -ina, spadá i předmět této práce, sufix -ák. Jako příklady univerbizace mohou posloužit: *angličtinář* (učitel angličtiny), *lanovka* (lanová dráha), *pracák* (úřad práce), *nákladák* (nákladní automobil), *tiskovka* (tisková konference) apod. Za projev univerbizační tendence lze považovat i vznik složenin z dvojslovných pojmenování, př. *gramodeska* (gramofonová deska), *elektrospotřebič* (elektrický spotřebič).

Jak je možné vidět z uvedených příkladů, univerbizace často zachovává tvar jednoho rodu, a sice rod původního substantiva z dvouslovného pojmenování: *obchodák* (obchodní dům), *blešák* (bleší trh), *tatarka* (tatarská omáčka), *sodovka* (sodová voda). Dochází však i k opačné situaci, kdy se rod výchozího substantiva mění: *občanka* (občanský průkaz) – z mužského na ženský; *Staromák* (Staroměstské náměstí), *svařák* (svažené víno), *hlavák* (hlavní nádraží) – ze středního na mužský. Kromě toho existují páry univerbizovaných substantiv mužského a ženského rodu, odvozených od stejného základu, přičemž každé z nich zachovává rod původního substantiva, například: *cestovka* (cestovní kancelář) – *cesták* (obchodní cestující). Vyskytují se i přechýlená slova: *angličtinář/ka*, *prvňák/prvňačka*. Z tohoto vyplývá, že ve většině případů adjektivum z dvouslovného pojmenování není substantivizováno v celém svém sémantickém spektru, nýbrž pouze v tom významu, který mělo ve spojení s příslušným substantivem. Diferenciace v rodu slouží k rozlišení významu, ale jen málo přispívá k odstranění případné nesrozumitelnosti, způsobené homonymií.

Homonymie vzniká v důsledku stírání původní motivace v průběhu univerbizace. Tak *cesták* označuje zároveň cestovní účet a obchodního cestujícího, *autobusák* může odkazovat jak na

autobusové nádraží, tak i na řidiče autobusu, *modráky* jsou buď modřiny nebo pracovní oděv (jednou je základem femininum, podruhé maskulinum, u pluralií tantum se povědomí o rodě stírá), a univerbum *škodovka* má dokonce tři významy: továrna, automobil a lokomotiva (Hauser, 1980, s. 25, 31). Jako příklad ze současného jazyka uveďme již zmíněné slovo *multák*, které může označovat multivitaminový nápoj i multikino. V prvním případě jde o čistou univerbizaci, v druhém o zkrácené slovo, které má v jazyce podobnou úspornou funkci jako univerbum.

Podle M. Dokulila (1962, s. 28) je univerbizace pojem z jiné roviny, než jsou slovotvorné způsoby (jako derivace a kompozice) a postupy (jako sufixace a změna tvaroslovné charakteristiky), a slouží k úpravě již hotového pojmenování pomocí způsobů a postupů ve vlastním smyslu slovotvorných. U Dokulila (1962, s. 115) najdeme následující definici univerbizace:

O univerbizaci jde tehdy a jen tehdy, když v jazyce koexistuje dvouslovné pojmenování, nikoli jen jako příležitostné vyčlenění zvláštního druhu uvnitř pojmenování rodového, nýbrž jako syntetické ustálené pojmenování ...

Abychom si ověřili, jedná-li se v daném konkrétním případě o univerbizaci, potřebujeme zjistit, koexistuje-li v jazyce dané slovo s odpovídajícím víceslovným pojmenováním. Například, slovo *obývánk* existuje současně a paralelně s více formálním pojmenováním *obývací pokoj* a tím pádem ho můžeme hodnotit jako univerbizaci.

Toto kritérium budeme používat při kategorizaci v rámci dané práce a budeme univerbizaci hledat především u slov hovorových a slangových, respektive u slov, která přešla ze slangu do jazyka spisovného a která jsou dosud pocíťována jako pojmenování zkrácená na pozadí víceslovného odborného názvu, a také u termínů (Dokulil, 1962, s. 117).

S univerbizací také souvisí další jevy vznikající v důsledku jazykové ekonomie, například denominativa z iniciálových zkratk nebo zkratkových slov motivovaných právě univerbizací. Sem patří velký počet slov běžně užívaných v mluveném jazyce: *festák* (festival), *noťas* (notebook), *sámoška* (samoobslužná prodejna), *klimoška* (klimatizace), *fitko* (fitness centrum), apod.

Kromě toho v moderním jazyce je velmi rozšířené tvoření jednoslovných pojmenování spojováním zkratk nebo zkratkových slov z cizích základů a domácích slovotvorných afixů:

ajták (z angl. zkratky „IT“), *sekáč* (z angl. „second hand“) anebo teď už zastaralé *vekslák* (z něm. „wechseln“) či dobový název *Enderák* (ze zkratky NDR – Německá demokratická republika).

3 PRAKTICKÁ ČÁST

V druhé, praktické části práce, představíme zpracování vyexcerpovaného materiálu. Pokusíme se o jeho kategorizaci na základě analýzy korpusových dat.

3.1 Polyfunkční sufix -ák v českých korpusech ČNK

Praktická část práce je věnována statistickému zpracování a analýze výskytů lemmat se sufixem -ák v korpusech českého jazyka ČNK. Během první etapy byl vytvořen výběr korpusových hesel, obsahujících zkoumaný sufix. Následně byly výsledky roztrženy na základě: (1) sémantického třídění, (2) slovotvorné kategorie a (3) typu odvozování, a poté provedené třídění bylo použito pro porovnání obsahu dvou použitých korpusů a interpretaci výsledků.

3.1.1 Pojem „lemma“

Pro praktickou část dané práce je také důležitý pojem „lemma“ čili „základní slovníkový tvar“. Tímto pojmem rozumíme základní podobu slova, která ve slovnících reprezentuje souhrn všech tvarů určité lexikální jednotky. Jako slovníkový tvar substantiv se v češtině zpravidla používá nominativ singuláru.

3.1.2 Použité nástroje

Statistické zpracování dat bylo provedeno na základě Českého národního korpusu, spravovaného Ústavem Českého národního korpusu a dostupného na stránkách korpus.cz. Data byla získána pomocí rozhraní KonText, které umožňuje vyhledávání v korpusech pomocí speciálních dotazů a další zpracování dat.

Jako zdroj jazykových dat jsem zvolil dva korpusy. Jako hlavní byl zvolen korpus *SYN2010*, synchronní reprezentativní korpus současné psané češtiny. Tento korpus obsahuje 100 milionů textových jednotek a je složen ze tří hlavních a poměrně vyvážených zdrojů: 40% z beletrie, 27% z odborné literatury a 33% z publicistiky. Korpus obsahuje odbornou

literaturu publikovanou po roce 1989 a publicistiku z let 2005 až 2009. Co se týká beletrie, v korpusu je sice zastoupena i ta starší, přesto největší podíl mají novější texty (Cvrček, 2015b).

Korpus *SYN2010* je lemmatizován a morfologicky označován. To znamená, že korpus prošel morfologickou analýzou, během níž každému slovu bylo přiřazeno jeho lemma (základní slovníkový tvar) a byly přiděleny morfologické údaje včetně slovního druhu (morfologické značky, tj. tagy). Takovéto zpracování korpusu umožňuje mimo jiné snadné vyhledávání konkrétního slovního druhu. Pro pokročilé vyhledávání v korpusech ČNK je používán speciální jazyk.

S ohledem na mluvený charakter sufixu -ák jsem jako vedlejší korpus pro výzkum také zvolil korpus mluveného jazyka *ORAL2013*. Tento korpus zachycuje spontánní konverzaci výhradně v neformálních komunikačních situacích. Korpus *ORAL2013* se skládá z 835 nahrávek z let 2008 až 2011; v sondách vystupuje celkem 2 544 mluvčích, z toho 1 297 unikátních. Korpus pokrývá celé území České republiky, zatímco celková délka nahrávek je téměř 300 hodin (Cvrček, 2015a).

Přestože korpus *ORAL2013* není lemmatizován ani morfologicky označován, což znamená, že nemáme možnost určit v dotazu slovní druh ani další morfologické kategorie, pro účely výzkumu v rámci této práce se nicméně podařilo tato omezení snadno obejít, jak je uvedeno níže v popisu použitého dotazu.

3.2 Pořizování a statistické zpracování výsledků

Morfologická analýza sufixu -ák je poměrně jednoduchá, poněvadž se tento formant vždy vyskytuje na konci slova a jediná změna v jeho struktuře, ke které dochází při jeho spojení s jinými sufixy a koncovkami, spočívá ve střídání písmen „k“ a „c“ před měkkou koncovkou „i“. Díky těmto vlastnostem vyhledávání substantiv se sufixem -ák je v korpusech poměrně snadným úkolem. V lemmatizovaných korpusech stačí v podstatě nalézt všechna slova, tj. lemmata končící na „ák“, zatímco v nelemmatizovaných korpusech je nutno do dotazu přímo zadat všechny varianty formantu -ák ve všech pádech a číslech (podrobněji viz níže).

3.2.1 Obecné charakteristiky korpusového dotazu a jeho struktura

Pro vyhledávání v korpusech pomocí rozhraní KonTex je používáno dotazů různého druhu – slovní tvar, lemma, základní, fráze, podřetězec a CQL (Corpus Query Language). Poslední z těchto typů je nejobtížnější a tak má největší použití pro profesionální cíle. Jeho základem je dotaz na jednu pozici (slovo) v korpusu:

```
[atribut="hodnota"]
```

kde atributem se myslí atribut poziční (word, lemma, tag apod.), hodnotou je pak samotný hledaný výraz, nebo vzor specifikovaný pomocí regulárního výrazu (Křen, 2016). Dotaz na více pozic současně (tj. na sekvenci slov) vzniká zřetěžením dotazů na jednotlivé pozice:

```
[atribut="hodnota"][atribut="hodnota"]
```

zatímco se atributy, které by se měly aplikovat ke stejnému heslu, spojují znakem „&“:

```
[atribut="hodnota" & atribut="hodnota"]
```

Pro dotazy, nutné v rámci této práce, bylo použito atributů „lemma“, „tag“ a „word“. Atribut „lemma“ určuje slovníkovou podobu slova, která při automatickém zpracování (lemmatizaci) je přidělována každé formě v korpusu (viz oddíl 3.1.1 Pojem „lemma“ výše). Atribut „tag“, čili morfologická značka, určuje gramatické parametry dotazu a je tvořen sekvencí písmen a čísel, které mají specifický význam podle toho, na jaké pozici v dotazu stojí. Například písmeno „N“ označuje substantivum, když stojí v první pozici a určuje slovní druh, v druhé pozici však označuje neutrum, tj. dodatečnou vlastnost již určeného slovního druhu. Atribut „word“ určuje jednotlivý slovní tvar a zachovává morfologickou (a případně i pravopisnou) podobu slova (na rozdíl od lemmatu, které reprezentuje množinu tvarů).

3.2.2 Dotazy pro vyhledávání v korpusech

Pro vyhledávání v korpusu *SYN2010* jsem použil následující dotaz:

```
[lemma=".{2,}ák" & tag="N.[MI].*"]
```

První položka tohoto dotazu (před znakem „&“) určuje vyhledávání všech slov, pro které jejich lemma (čili základní slovníkový tvar) končí příponou -ák, přičemž zároveň vyřazuje

výskyty skládající se ze tří písmen, poněvadž u nich prvek „ák“ nikdy není příponou, ale částí slovního základu. Jsou to především slova: *žák, mák, hák* a *lák*. Druhá položka (za znakem „&“) omezuje výskyty jen na nominativa mužského rodu životného a neživotného, čímž se dosahuje vymezení některých příslovcí, částic a jiných slovních druhů končících řetězcem znaků „ák“ (zejména: *kdepák, ňák, tááák* apod.).

Vzhledem k tomu, že korpus *ORAL2013* není lemmatizován, pro vyhledávání v něm bylo zapotřebí použít jiný dotaz:

```
[word=".{2,}(ák|áka|áku|ákovi|ákem|áci|áků|ákům|áky|ácích)"]
```

V položce jsem musel explicitně vyjmenovat všechna možná zakončení substantiv se sufiksem -ák v různých pádech a v obou číslech.

3.2.3 Frekvenční distribuce

Kompletní výsledek hledání v korpusech *SYN2010* a *ORAL2013* je uveden v přílohách 1 a 2.

Výskyty v nich uvedené jsou uspořádány podle frekvenční distribuce, tj. podle počtu jednotlivých výskytů v korpusu. Frekvenční distribuce je jednou z funkcí, které poskytuje rozhraní KonText. Kvůli velkému počtu primárních výskytů jsem ji použil, abych vyčlenil nejvíce frekventovaná substantiva s příponou -ák a mohl na ně zaměřit svůj výzkum (viz oddíl 3.2.4 *Úprava výsledků* níže).

3.2.4 Úprava výsledků

Vzhledem k tomu, že dále chceme pracovat pouze se substantivy odvozenými pomocí suffixu -ák a dosáhnout pokud možno statisticky srozumitelných výsledků, bylo nutné ze získaných seznamů odstranit výskyty, které tento požadavek nesplňují.

Za prvé ze seznamu kvůli objemu materiálu byla odstraněna substantiva s jediným výskytem. Taková substantiva tvoří přibližně třetinu všech výskytů v korpusu *SYN2010* a kolem 60% výskytů v korpusu *ORAL2013*. Taková slova můžeme zařadit do kategorie *hapax legomena*. Tímto pojmem označujeme slova, která se v určitém textovém souboru vyskytují jen jednou.

Dle manuálu ČNK, jejich využití v lingvistickém výzkumu je dosud nejasné a sporné, protože jediný výskyt neumožňuje zobecňovat nebo vytvářet jakékoli závěry o vlastnostech a způsobech užití té či oné jednotky (Křen, 2010). Zároveň se ale mezi těmito slovy, jak už bylo zmíněno výše, mohou vyskytovat neologismy, což je v případě sufixu -ák vzhledem k jeho produktivitě údaj velmi cenný.

Mezi slovy s jediným výskytem najdeme jak slova poměrně známá, například: *nepiják* (člověk, který nepije), *gruziňák* (gruzínský koňák), *včelák* (druh hmyzu), *prsták* (kdo má prsty), *bedňák* (v moderním jazyce – asistent zvukaře na koncertě), *dvojkolák* (vozík o dvou kolech), *páteřák* (chránič na páteř), *texták* (textový editor), *sokolovák* (člen Sokola), *otevírák* (nástroj určený k otevírání), *oblbovák* (narkóza), *jídelňák* (jídelní lístek), *medák* (čmelák), *cedák* (cedník), *vyšetřovák* (vyšetřovatel), *brabčák* (vrabec), *devizák* (devizový příslib), tak i slova málo rozšířená, anebo i taková, jejichž význam nejde snadno odhadnout mimo kontext, například: *portulák*, *waltrovák*, *uběrák*, *zrnovák*, *hliňák*, *tralalák*, *střihák*, *zabalák*, *luďák*, *hapták*, *sádelňák*, *profák*, *naťák*, *syrák*, *buršák* apod.

Dálo by se říct, že by pro účely zkoumání tak živého a produktivního prostředku, jakým je sufix -ák, analýza hapaxů mohla být přínosná, poněvadž by se právě mezi nimi mohly nalézt neologismy a autorské novotvary. Nicméně vzhledem k vysoké vazbě substantiv tvořených pomocí sufixu -ák na kontext, třídění jejich významů vně kontextu se jeví jako potenciálně nespolehlivé. Dle výzkumu R. Novotné (2012), sufix -ák zaujímá druhou pozici mezi sufixy nejvíce charakteristickými pro hapax legomena v korpusu *SYN* (jsou to: -ovský, **-ák**, -ista, -ing/-ink, -ovitý). Jako příklady hapaxů se sufixem -ák Novotná uvádí *mirindák* (utvořeno od nápoje Mirinda), *medvěďák* (několikadenní spánek), *odkladák* (domov důchodců), *klafák* (velký zub) a další. Zároveň se v její práci potvrzuje, že „některé hapaxy jsou sémanticky nejasné, protože jediný kontext nemůže poskytnout dostatečnou informaci o významu daného lexému“ (Novotná, 2012). Dalším nesmírně zajímavým faktem je, že na rozdíl od frekventovanějších slov a hlavně transparentní motivace pojmenování při tvorbě univerb, u citovaných hapaxů často jde o přenesený, asociativní či metaforický význam (srov. *medvěďák*, *odkladák*). Toto jen potvrzuje výše zmíněnou hravost a autorskou tvořivost, které doprovázejí vznik takovýchto pojmenování.

Za druhé ze seznamu byla odstraněna substantiva, u kterých zakončení na „ák“ neobsahuje zrovna sufix -ák, nýbrž jde o část slovního základu, a tím pádem se nejedná o slova odvozená. Z nejvíce frekventovaných výskytů do této kategorie spadají: *žák*, *pták*, *hák*, *lák*, *mák*, *tabák*,

barák, čmelák, luňák, pašák, slimák, maják. Jak již bylo zmíněno, třípísmenná slova a jiné slovní druhy kromě substantiv mužského rodu byly odstraněny již v etapě zadání dotazu. Při úpravě výsledku hledání v korpusu *ORAL2013* také byla odstraněna některá slova, zahrnující řetězce znaků, uvedené v dotazu, které však nejsou alomorfy formantu -ák. Jde především o slovní variantu „práci“ (obsahující řetězec „áci“) apod. Slova vyjmenovaná výše tvoří cca 13% z celkového počtu výskytů. Slova odstraněná automaticky při podání dotazu nejsou zahrnuta do tabulek, a tak jejich počet nelze zjistit.

Za třetí ze seznamu byla odstraněna vlastní jména, jakožto: Novák, Dvořák, Čermák, Kozák, Husák, Novák, Horák, Jandák, Svěrák, Vlasák, Izák, Sedlák a spousta dalších. Příjmení končící na -ák tvoří jen malou část původních výskytů (cca 1%).

Kromě toho, poněvadž korpus *ORAL2013* není lematizován, výsledky získané z něj musely být manuálně uspořádány podle lemat. Toto bylo nutné za účelem provedení správné statistické analýzy a porovnání s výsledky získanými z korpusu *SYN2010*. Například, výskyty: *bramborák* (2), *bramboráka* (2), *bramboráku* (1), *bramboráků* (1) a *bramboráky* (13) byly sloučeny pod jednotlivým lematem *bramborák* s celkovým počtem 19 výskytů. Uvedeným způsobem byly ručně zpracovány všechny výsledky z korpusu *ORAL2013*. Posléze byly výskyty ve tvaru odlišném od lematu převedeny na formu jednotného čísla nominativu.

3.2.5 Přehled nejvíce frekventovaných slov

Je nutno říci, že se výsledky, pořízené ze dvou korpusů, ve značné míře nekryjí. Korpus *SYN2010* obsahuje jen cca 30% výsledků nalezených v *ORAL2013*, zatímco *ORAL2013* obsahuje jen cca 50% výsledků nalezených v *SYN2010*.

Níže uvedená tabulka znázorňuje 100 nejfrekventovanějších výskytů z obou korpusů a zároveň ilustruje, nakolik se oba dva korpusy navzájem kryjí. Políčka bez barvy jsou přítomna v první stovce obou korpusů, zatímco barevně označená políčka ukazují na to, že daný výskyt najdeme v první stovce pouze v jednom z korpusů:

	SYN2010	ORAL2013
1.	divák	chudák
2.	voják	foťák
3.	zpěvák	voják
4.	Polák	panák
5.	chudák	prvák
6.	Slovák	řidičák
7.	rodák	spacák
8.	školák	Slovák
9.	hlupák	mražák
10.	leták	Polák
11.	sedlák	leták
12.	zabiják	třeťák
13.	kuřák	Rusák
14.	zobák	druhák
15.	panák	mrazák
16.	padák	Pražák
17.	předák	divák
18.	parták	zpěvák
19.	žebrák	kilák
20.	Klusák	čumák
21.	vysokoškolák	sedlák
22.	liják	tepláky
23.	špičák	štvrták
24.	řidičák	čurák
25.	tulák	šroubovák
26.	Hanák	činžák
27.	darebák	květák
28.	věšák	držák
29.	ukazovák	fešák
30.	bouřlivák	centák
31.	činžák	vrták
32.	držák	Brňák
33.	vodák	naviják
34.	trhák	Rakušák
35.	pěšák	tuňák
36.	pytlák	basák
37.	tučňák	kuřák
38.	spacák	širák
39.	středoškolák	vysokoškolák

40.	strašák	ajták
41.	tuňák	pančák
42.	čumák	pařák
43.	ubožák	zobák
44.	mrzák	věšák
45.	estébák	žebrák
46.	foťák	Bosák
47.	ovčák	křoviňák
48.	mrazák	Moravák
49.	Pražák	reprák
50.	pašerák	chytrák
51.	tahák	feťák
52.	kozák	puberták
53.	blondák	saniták
54.	auták	táborák
55.	zadák	vlčák
56.	pasák	Zelňák
57.	divočák	blondák
58.	vrták	hasičák
59.	dobrák	měšťák
60.	pošťák	ódéesak
61.	tesák	pošťák
62.	klusák	rolák
63.	esesák	rypák
64.	květák	sušák
65.	feťák	sympaťák
66.	větrák	tahák
67.	ležák	větrák
68.	puberták	vodák
69.	Čechoslovák	votvírák
70.	sedák	zeťák
71.	gestapák	deváták
72.	piják	lesák
73.	fešák	lyžák
74.	drsnák	nervák
75.	vlčák	padák
76.	hořák	chcípák
77.	nekuřák	dobrák
78.	naviják	estébák
79.	křížák	koumák

80.	plovák	koupák
81.	předškolák	otvírák
82.	silák	servisák
83.	všivák	slejšák
84.	chytrák	sporták
85.	táborák	zabiják
86.	smeták	bavorák
87.	kluzák	čmelák
88.	plameňák	dospělák
89.	ušák	elasták
90.	levák	fóliák
91.	šroubovák	hořák
92.	junák	kamionák
93.	čtvrťák	Křížák
94.	čmelák	ležák
95.	širák	magneták
96.	levičák	parták
97.	rolák	plasták
98.	bzučák	plovák
99.	mazák	podpantoflák
100.	lesák	prasečák

Tabulka 1: Porovnání 100 nejfrekventovanějších výskytů v korpusech SYN2010 a ORAL2013

Tato data znázorňují, jak výrazně se liší výskyty substantiv se sufixem -ák ve dvou použitých korpusech. Tak velký rozdíl by se dal vysvětlit velkou produktivností a mluveným rázem sufixu -ák, kvůli nimž jeho využití nemá výrazný ustálený charakter, čímž vzniká poměrně velký objem slovní zásoby, která se v různých jazykových vrstvách neopakuje. Příklady z mluveného korpusu nedoložené v SYN2010 jsou zároveň ilustrací nejfrekventovanějších slov v mluveném jazyce zakončených sufixem -ak: např. *prvák, druhák, třeták, štvrták, Rusák, čurák, ajťák, pařák, sympaták, koumák, sporták, dospělák, odéesák, měšťák, zeťák, chcípák, slejšák, sporták, bavorák, nervák, podpantoflák*.

3.2.6 Třídění výsledků

Jak je možné vidět v přílohách 1 a 2, po úpravě byly výsledky pro účely třídění manuálně označeny písmeny Č – člověk, Z – zvíře, V – věc, M – místo, A – abstraktum pro třídění sémantické; písmeny ČIN – činitelské, KON – konatelské, VL – vlastnosti, OBYV – obyvatelské, PŘÍ – příslušenské, PRO – prostředků, AUG – augmentativa, PŘECH – přechýlená jména pro třídění podle slovotvorných kategorií; a nakonec písmeny DV – deverbativa, DS – desubstantiva, DA – deadjektiva, DN – denumeralia a U – univerbizace pro třídění podle typu odvozování. Tato označení pak byla použita pro tvoření statistických diagramů, které jsou uvedeny a diskutovány v příslušných podkapitolách níže.

V několika málo případech, kdy to či ono slovo nemohlo být jednoznačně zařazeno do určité kategorie (například, *cesták*, *druhák*, *šmírák* – poněvadž může označovat jak člověka, tak i věc), za účelem usnadnění práce byl použit jeden z významů, což by statisticky nemělo mít podstatný vliv na celkový výsledek.

Vykřičníkem v tabulkách jsou označena substantiva odvozená od problematického základového slova nebo cizojazyčného základu, rovněž také i substantiva, u kterých zakončení na „ák“ neobsahuje zrovna sufix -ák (podrobněji viz výše v oddílu 3.2.4 *Úprava výsledků*).

3.3 Prezentace výsledků – třídění sémantické (člověk, zvíře, věc, místo, abstraktum)

Sémantické třídění substantiv se sufixem -ák pro každý ze dvou zkoumaných korpusů je představeno v diagramu 1 a tabulce 2.

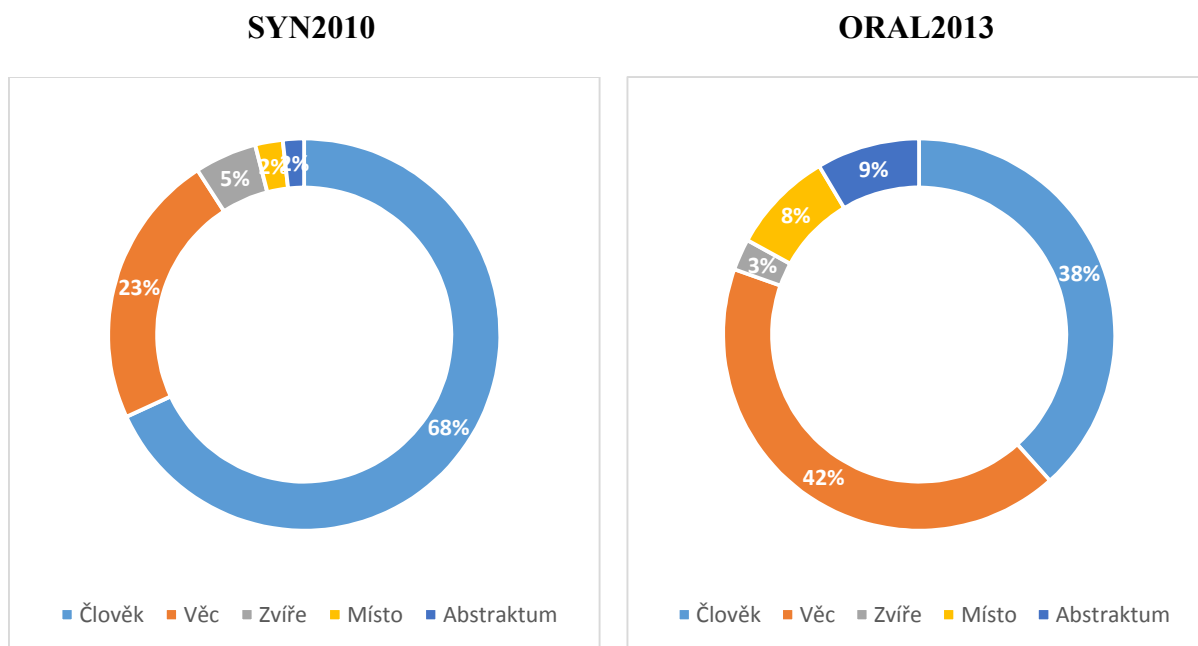


Diagram 1: Pokrytí substantiv patřících k jednotlivým sémantickým kategoriím

	Člověk	Věc	Zvíře	Místo	Abstraktum	<i>Celkem</i>
SYN 2010	72 354	24 142	5 463	2 418	1 860	92 478
	68%	23%	5%	2%	2%	100%
ORAL 2013	1 302	1 426	88	288	289	3 393
	38%	42%	3%	8%	9%	100%

Tabulka 2: Počty lemmat v jednotlivých sémantických kategoriích

Celková sémantická distribuce v obou korpusech vykazuje jisté rozdíly a odpovídá očekávané distribuci vycházející z empirické jazykové zkušenosti. Největší počet výskytů vidíme v kategoriích *člověk* a *věc*, zatímco kategorie *zvíře*, *místo* a *abstraktum* mají vedlejší pozice. Nicméně v psaném korpusu *SYN2010* pozorujeme velkou (skoro dvojnásobnou) převahu pojmenování lidí v porovnání s korpusem *ORAL2013*. Tento fakt svědčí o větší spisovnosti daného typu substantiv. Tato převaha je dána nejen rozdílem ve složení dvou korpusů, ale také odlišnou frekvenční váhou jednotlivých slov, jež se vyskytují v obou korpusech. Jako příklad daného jevu můžeme uvést slovo *divák*, které se v korpusu *SYN2010* ujímá první pozice s velkou převahou (spolu se substantivem *voják*) nad ostatními výskyty, zatímco se v korpusu *ORAL2013* slovo *divák* nachází až na devatenácté pozici. Naproti tomu v mluveném korpusu se dvakrát častěji než v psaném korpusu vyskytují lemmata značící *věc* a čtyřikrát častěji lemmata označující *místo* a *abstrakta*. Jedná se zřejmě o frekventovaná

univerbizační označení náměstí, nádraží, sídliště, mostů apod. typu *Václavák*, *Staromák*, *hlavák*, *Čerňák*, *Nuselák*, *Libeňák* apod.

3.4 Presentace výsledků – třídění podle slovotvorných kategorií (názvy obyvatelské, vlastnosti, prostředků apod.)

Slovotvorné třídění substantiv se sufixem -ák pro každý ze dvou zkoumaných korpusů je představeno v diagramu 2 a tabulce 3.

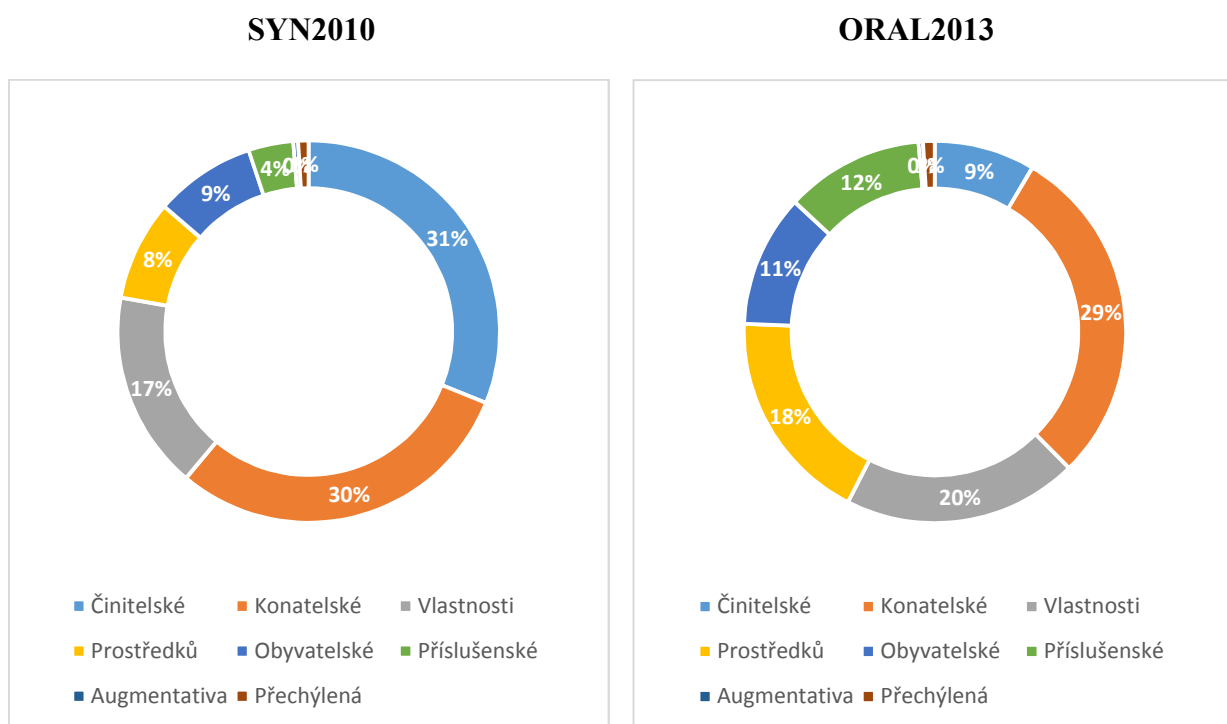


Diagram 2: Pokrytí substantiv patřících k jednotlivým slovotvorným kategoriím

	Činitel- ské	Kona- telské	Vlastno- sti	Pro- středků	Obyva- telské	Příslu- šenské	Aug- men- tativa	Přechý- lená	<i>Celkem</i>
SYN 2010	29 838	28 806	16 088	8 169	8 192	3 639	402	848	95 982
	31%	30%	17%	8%	9%	4%	<1%	<1%	100%
ORAL 2013	211	719	492	449	277	290	9	25	2 261
	9%	30%	20%	18%	11%	12%	<1%	<1%	100%

Tabulka 3: Počty lemmat v jednotlivých slovotvorných kategoriích

Distribuce výskytů v psaném a mluveném korpusu má v mnoha slovotvorných kategoriích shodný charakter. Výjimkou je však převaha substantiv patřících do činitelské třídy v korpusu *SYN2010*. Z tohoto můžeme vyvodit závěr, že jsou substantiva činitelského typu typická pro spisovný jazyk. Jsou to deverbativa se sémantickým významem *člověk* a *věc*, v první pětce činitelských jmen v korpusu *SYN2010* však najdeme jen typ *člověk*: *divák*, *zpěvák*, *zabiják*, *kuřák*, *žebrák*. Naopak, třída prostředků a třída příslušenská mají mnohem vyšší výskyt v mluveném korpusu: třída prostředků se vyskytuje víc než dvakrát častěji, zatímco třída příslušenská až třikrát častěji v korpusu *ORAL 2013*. Toto poukazuje na větší použití substantiv těchto tříd v hovorovém jazyce. Nejfrekventovanějšími výskyty jmen prostředků a jmen příslušenských v korpusu *ORAL 2013* jsou: *foťák*, *prvák*, *spacák*, *leták*, *třeťák*.

3.5 Presentace výsledků – třídění podle typu odvozování (deverbativa, desubstantiva, deadjektiva apod.)

Třídění substantiv se sufixem -ák podle typu odvozování pro každý ze dvou zkoumaných korpusů je představeno v diagramu 3 a tabulce 4.

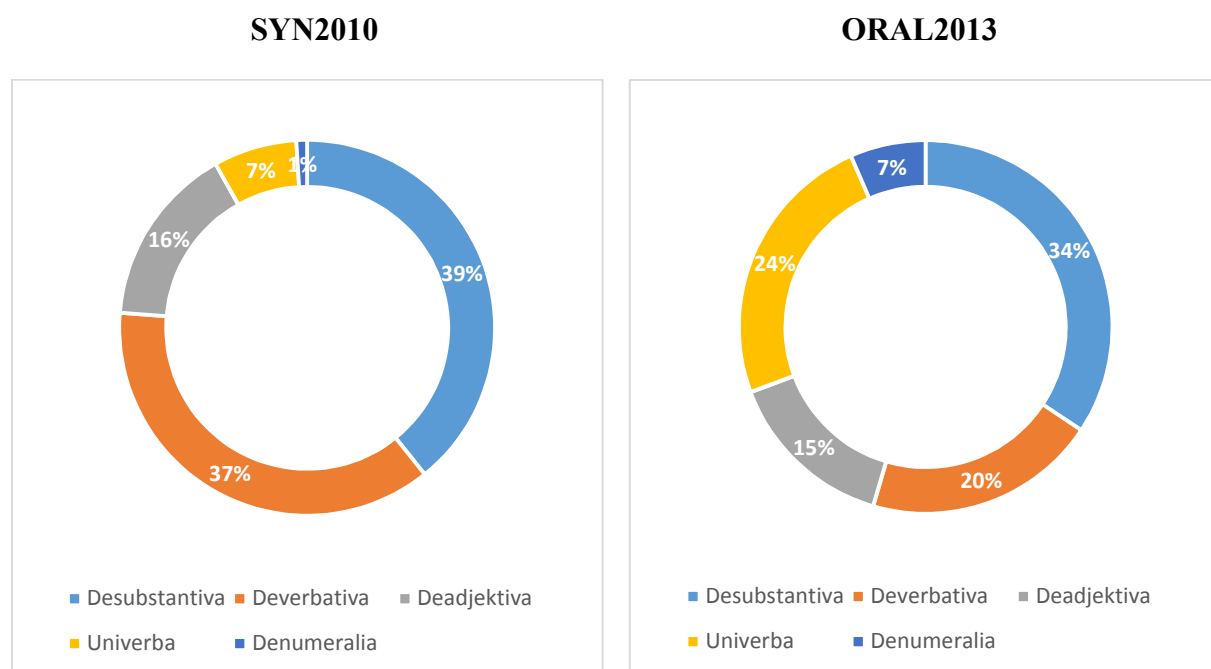


Diagram 3: Pokrytí substantiv různého typu odvozování

	Desubstan- tiva	Deverbativa	Deadjektiva	Univerba	Denum- ralia	<i>Celkem</i>
SYN 2010	40 704	38 445	16 189	7 452	952	103 742
	39%	37%	16%	7%	1%	100%
ORAL 2013	1 120	660	483	787	214	3 264
	34%	20%	15%	24%	7%	100%

Tabulka 4: Počty lemmat různého typu odvozování

Výsledky třídění podle typu odvozování ve dvou korpusech můžeme charakterizovat jako podobné u *desubstantiv* a *deadjektiv*. Přesto se tato klasifikace také vyznačuje třemi výraznými odchylkami, a to u typu *deverbativ* (skoro dvojnásobná převaha v *SYN2010*), *univerb* a *denumerálii* (ještě výraznější převaha v *ORAL2013*).

První odlišnost poukazuje na spisovný charakter deverbativního typu odvozování.

Druhá odlišnost je spojená s jevem univerbizace, charakteristickým pro český jazyk, a prokazuje, že se při univerbizaci ve značné míře využívá mluvená vrstva jazyka. Je to kompaktní a oblíbený jazykový prostředek, mnohem více rozšířený v mluveném projevu kvůli své emotivní zbarvenosti a ekonomičnosti. Totéž lze říci o denumeráliích jako *prvák*, *druhák*, *třeták*, která se hojně vyskytují v mluveném jazyce.

3.6 Prezentace výsledků – statistická distribuce mezi slovotvornými a sémantickými kategoriemi

Dalším cílem výzkumu bylo zjištění statistických souvislostí mezi několika kategoriemi. Jak již bylo zmíněno v oddílu 2.5 *Třídění podle typu odvozování*, každá slovotvorná kategorie je většinou vázána na jeden určitý typ odvozování a tím pádem by porovnání slovotvorných kategorií a typů odvozování pravděpodobně nepřineslo žádné zajímavé výsledky. Proto jsem se rozhodl pro zkoumání statistických zákonitostí mezi slovotvornými a sémantickými kategoriemi, totiž kategoriemi, které nejsou na sebe tak pevně vázány. Výsledky tohoto výzkumu jsou uvedeny v diagramu 4 a tabulce 5.

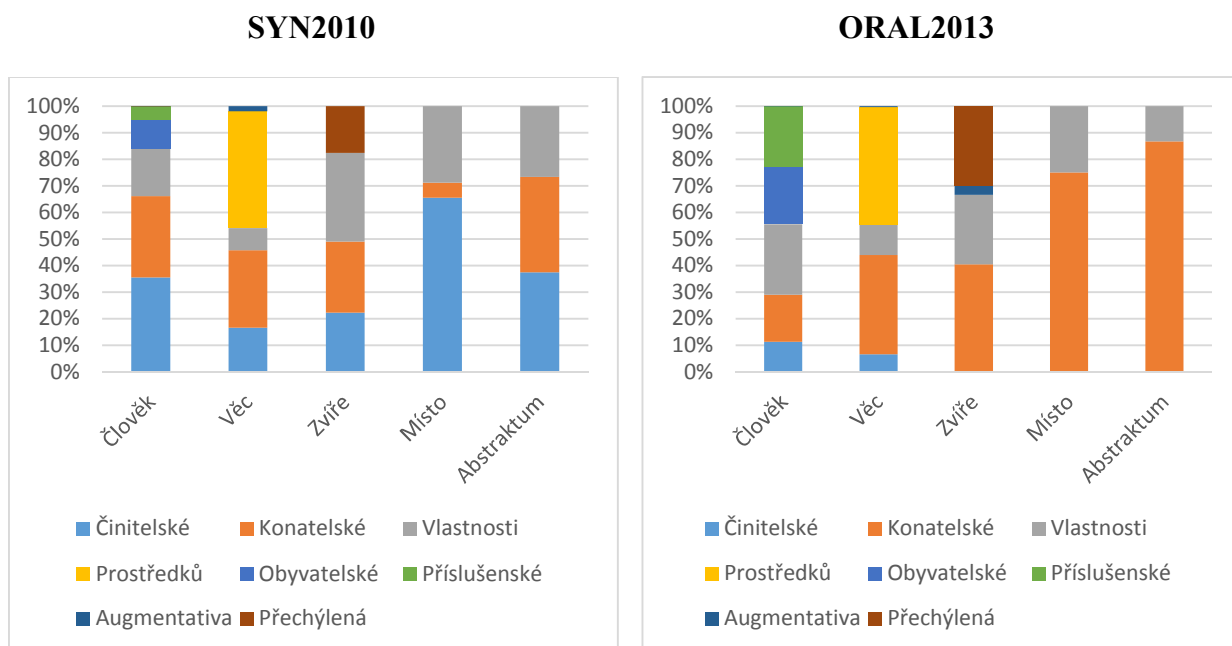


Diagram 4: Distribuce slovotvorných kategorií v každé ze sémantických kategorií

SYN2010

	Činitel- ské	Kona- telské	Vlastno- sti	Pro- středků	Obyva- telské	Příslu- šenské	Aug- men- tativa	Přechý- lená	<i>Celkem</i>
Člo- věk	25 431	21 921	12 787	0	7 734	3 639	64	9	71 585
	36%	31%	18%	0%	11%	5%	0%	0%	100%
Věc	3 088	5 418	1 539	8 192	0	0	338	0	18 575
	17%	29%	8%	44%	0%	0%	2%	0%	100%
Zvíře	1 067	1 275	1 599	0	0	0	0	839	4 780
	22%	27%	33%	0%	0%	0%	0%	18%	100%
Místo	57	5	25	0	0	0	0	0	87
	66%	6%	29%	0%	0%	0%	0%	0%	100%
Abs- trak- tum	195	187	138	0	0	0	0	0	520
	38%	36%	26%	0%	0%	0%	0%	0%	100%

ORAL2013

	Činitel- ské	Kona- telské	Vlastno- sti	Pro- středků	Obyva- telské	Příslu- šenské	Aug- men- tativa	Přechý- lená	<i>Celkem</i>
Člo- věk	144	227	339	0	277	290	2	0	<i>1 279</i>
	11%	18%	27%	0%	22%	23%	0%	0%	<i>100%</i>
Věc	67	380	118	449	0	0	4	0	<i>1 018</i>
	7%	37%	12%	44%	0%	0%	0%	0%	<i>100%</i>
Zvíře	0	34	22	0	0	0	3	25	<i>84</i>
	0%	40%	26%	0%	0%	0%	4%	30%	<i>100%</i>
Místo	0	6	2	0	0	0	0	0	<i>8</i>
	0%	75%	25%	0%	0%	0%	0%	0%	<i>100%</i>
Abs- trak- tum	0	72	11	0	0	0	0	0	<i>83</i>
	0%	87%	13%	0%	0%	0%	0%	0%	<i>100%</i>

Tabulka 5: Distribuční počty lemmat slootovorných kategorií v každé ze sémantických kategorií

Dle očekávání celkově převládají substantiva činitelská, konatelská a nositelů vlastností, jsou to zdaleka nejrozšířenější typy. Substantiva některých typů se vyskytují jen v určitých sémantických kategoriích. Substantiva obyvatelská a příslušenská nalezneme jen v kategorii *člověk*, rovněž jako substantiva prostředků jen v kategorii *věc*.

Distribuce jednotlivých kategorií v korpusech *SYN2010* a *ORAL2013* vykazují do značné míry odlišné hodnoty, jež mohou souviset s různými typy korpusů. V tabulce 5 jsou tučně označeny nejčtenější výskyty. Některé z nich by se daly interpretovat nebo vysvětlit. Například převaha jmen činitelských a konatelských v kategorii *člověk* v korpusu *SYN2010* pravděpodobně souvisí s celkovou převahou jak dané kategorie, tak i daných slootovorných tříd v tomto korpusu (viz oddíly 3.3 *Prezentace výsledků – třídění sémantické* a 3.4 *Prezentace výsledků – třídění podle slootovorných kategorií* výše). Stejná procenta u jmen prostředků a v kategorii *věc* v obou korpusech znázorňují typický charakter této kategorie pro danou slootovornou třídu, univerzální pro mluvený i psaný jazyk.

Další odlišností je nepřítomnost substantiv činitelského typu v kategoriích *zvíře*, *místo* a *abstraktum* v korpusu *ORAL2013*. Když se podíváme do korpusu *SYN2010*, zjistíme, že jsou této kategorie ve skutečnosti v něm představeny jen několika slovy, která jsou ale poměrně frekventovaná: *klusák* a *plzák* (*zvíře*), *klouzák* a *koupák* (*místo*) a *stoják*, *zlepšovák*, *ploužák*, *ječák*, *dupák*, *lomcovák* a *rejdvák* (*abstraktum*). Tím pádem, motivy této odlišnosti jsou pravděpodobně jen statistické a jsou zapříčiněny poměrně malým objemem korpusu *ORAL2013*. Tak běžně užívaná slova jako *stoják*, *zlepšovák* nebo *ploužák* by se především měla vyskytnout právě v mluveném korpusu. Jejich absence souvisí s mírou reprezentativnosti, která u korpusu menšího rozsahu nemusí vždy být dostatečně zastoupena.

3.7 Prezentace výsledků – distribuce univerbizace v sémantických kategoriích

Kromě výše uvedených výsledků, zajímala mě také distribuce univerbizace jako jevu velmi specifického, a tak jsem prozkoumal její frekvenci mezi sémantickými kategoriemi. Získané výsledky jsou uvedeny v diagramu 5 a tabulce 6.

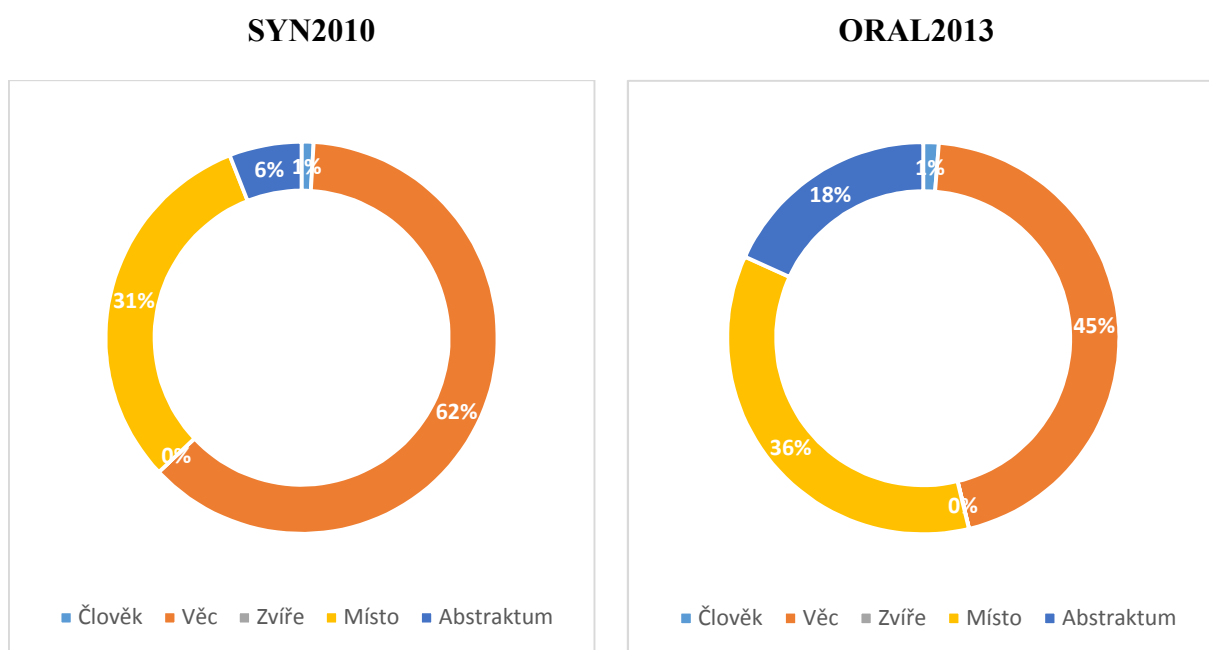


Diagram 5: Distribuce univerbizace v sémantických kategoriích

	Člověk	Věc	Zvíře	Místo	Abstraktum	<i>Celkem</i>
SYN 2010	77	4 604	5	2 301	465	7 452
	1%	62%	0%	31%	6%	100%

ORAL 2013	10	354	0	280	143	787
	1%	45%	0%	36%	18%	100%

Tabulka 6: Distribuční počty univerbizace v sémantických kategoriích

V diagramech pozorujeme velkou shodu mezi dvěma korpusem, což umožňuje udělat závěr o vysoké spolehlivosti zjištěné zákonitosti. Dvě odlišnosti spočívají ve trojnásobné převaze *abstrakt* v korpusu *ORAL2013* a o třetinu vyšším zastoupení *univerb* v pojmenování věcí v *SYN2010*. Korpus *ORAL2013* obsahuje celou řadu univerb, jež v korpusu *SYN2010* nenalezneme. Korpusy se liší hlavně abstrakty s malou frekvencí: *rodičák*, *absták*, *bercák*, *jarňák*, *jednodeňák*, *orienták*, *penziják*, *zápočták*, *cyklisták*. Jak je možné postřehnout, mnoho z těchto slov je odvozeno z víceslovných pojmenování, označujících aktivity různého druhu, jakožto tábory a srazy, finanční prostředky, mentální stavy apod. Sestava těchto slov svědčí o produktivnosti univerbizačního postupu v mluveném jazyce, kde se univerbizace používá pro zkrácené pojmenování složených, avšak frekventovaných záležitostí týkajících se každodenního života.

3.8 Zajímavý fakt

Pro úplnost bych také uvedl jedno zajímavé zjištění, které jsem vypožadoval během výzkumu, a sice fakt, že se kvůli své výstižnosti substantiva se sufixem -ák často vyskytují jako názvy filmů, jak českých, tak i jako překlady názvů zahraničních. Jako příklad uvedu *Zeták*, *Rakeťák*, *Chechták*, přičemž celkový počet jednoslovných filmových titulů v získaných výsledcích přesahuje desítku.

4 ZÁVĚR

V bakalářské práci jsme věnovali pozornost českému sufixu -ák a zpracování jeho distribuce na základě korpusového materiálu. Popsali jsme jeho vlastnosti jako slovotvorného prostředku, včetně takových charakteristik jako homonymita a produktivita. Byly probrány otázky citového zabarvení sufixu -ák, jakož i možné důvody jeho expresivity.

V rámci praktické části práce byly vytvořeny seznamy výskytu substantiv se sufixem -ák v korpusech českého jazyka *SYN2010* a *ORAL2013*, poskytovaných Českým národním korpusem. Seznamy byly získány pomocí standardních nástrojů rozhraní KonText, analýza významů a jejich klasifikace byla provedena převážně manuálně.

Hlavní výsledky jsou obsaženy v diagramech 1, 2 a 3 a tabulkách 2, 3 a 4, jež ukazují, jaké sémantické vlastnosti, slovotvorné charakteristiky a typy odvozování nese sufix -ák při praktickém použití v rámci současného jazyka a s jakými četnostmi. Nejvíce zastoupenými sémantickými kategoriemi jsou *člověk* a *věc*, nejvíce zastoupenými slovotvornými třídami jsou jména *činitelská*, *konatelská* a *vlastnosti*, nejvíce zastoupenými typy odvozování jsou *desubstantiva* a *deverbativa*.

Rozdíly mezi psaným a mluveným korpusem jsou přitom značné: v psaném korpusu se sémantická kategorie *člověk* vyskytuje dvakrát častěji než v mluveném, kdežto v mluveném korpusu se dvakrát častěji než v psaném vyskytují lemmata značící *věc* a čtyřikrát častěji lemmata označující *místo* a *abstrakta*. Pokud jde o výskyt lemmat v rámci jednotlivých slovotvorných kategorií, v psaném korpusu se třikrát častěji vyskytuje kategorie *činitelská*, zatímco v mluveném korpusu dvakrát častěji kategorie *prostředků* a třikrát častěji kategorie *příslušenská*. Z hlediska typu odvozování se pouze *desubstantiva* vyskytují stejně často jak v psaném, tak v mluveném korpusu. *Deverbativa* se vyskytují dvakrát častěji v psaném korpusu, zatímco se *univerba* a *denumeralia* vyskytují mnohem častěji v mluveném korpusu: *univerba* 3,5krát častěji, *denumeralia* až sedmkrát častěji.

Další závěry byly vypracovány na základě několika prvotních statistických výsledků a jsou obsaženy v diagramech 4 a 5 a tabulkách 5 a 6. Tyto výsledky zachycují statistickou distribuci mezi slovotvornými a sémantickými kategoriemi a distribuci univerbace v sémantických kategoriích. V obou korpusech distribuce má celkově rozdílný charakter. Nejvíce patrným rozdílem je nepřítomnost substantiv činitelského typu v kategoriích *zvíře*, *místo*, *abstraktum* a velká převaha *univerb* abstraktního typu v korpusu *ORAL2013*.

Výzkum sice v některých případech potvrdil shodnou tendenci v distribuci substantiv tvořených suffixem -ák v psaném a mluveném korpusu, avšak při různém rozložení jednotlivých sémantických typů, slovotvorných kategorií a typů odvozování. Popsané odchylky mohou do jisté míry souviset s nejednotným rozsahem obou korpusů, nicméně vypovídají též o odlišném charakteru psaného a mluveného jazyka.

5 SEZNAM POUŽITÉ LITERATURY

Primární

1. Cvrček, V. (2015a): *Korpus neformální mluvené češtiny ORAL2013*. Dostupné z: <http://wiki.korpus.cz/doku.php/cnk:ORAL2013>
2. Cvrček, V. (2015b): *Korpus SYN2010*. Dostupné z: <http://wiki.korpus.cz/doku.php/cnk:SYN2010>
3. Cvrček, V. (2013): *Hapax (legomenon)*. Dostupné z: [http://wiki.korpus.cz/doku.php/pojmy:hapax?s\[\]=hapax](http://wiki.korpus.cz/doku.php/pojmy:hapax?s[]=hapax)
4. David, M. (1999): *Tvoření slov v češtině*. České Budějovice: Pedagogická fakulta JU
5. Dokulil, M. (1967): *Tvoření slov v češtině*, díl 2. Praha: Academia
6. Dokulil, M. (1962): *Tvoření slov v češtině*, díl 1. Praha: Academia
7. Hauser, P. (1980): *Nauka o slovní zásobě*. Praha: SPN
8. Křen, M. (2016): *Dotazovací jazyk*. Dostupné z: https://wiki.korpus.cz/doku.php/pojmy:dotazovací_jazyk
9. Mitter, P. [b.r.]: *Lexikologie českého jazyka*. Ústí nad Labem: Studijní opora Katedry bohemistiky Pedagogické fakulty UJEP. Dostupné z: <http://www.pf.ujep.cz/pf/osobni-dokumenty/kbo/286-kbo-mitter-studijni-opora-lexikologie/file>
10. Novotná, R. (2012): Jazyková potencialita: studium na bázi hapaxů legomenon. In: *Gramatika & korpus. Gaudeamus*, Univerzita Hradec Králové, 1-8. Dostupné z: http://utkl.ff.cuni.cz/~rosen/public/GC2012/Konferencni_prispevky/NovotnaRenata.pdf

Sekundární

11. Čechová, M. a kol. (2008): *Současná stylistika*. Praha: NLN
12. Čermák, F. (2012): *Morfematika a slovotvorba češtiny*. Praha: NLN
13. Daneš – Dokulil – Kruhařa kol. (1967): *Tvoření slov v češtině 2: Odvozování podstatných jmen*. Praha: Academia
14. Helcl, M. (1963): *Univerbizace a její podíl při růstu dnešní slovní zásoby*. Praha: Academia
15. Janovec, L. (2007): K projevům jazykových vývojových tendencí v současné češtině. In: *NŘ*, roč. 90, 2007, č. 2
16. Jedlička, A. (1969): Univerbizace a multiverbizace v pojmenovacích strukturách. In: *AUC. Slavica Pragensia XI*. Praha: Karolinum, s. 93–101.

17. Kolářová, I. (2006): Konkurence univerbizovaných a víceslovných pojmenování v textech Českého národního korpusu. In: *Český jazyk a literatura v interakci – Język i literatura czeska w interakcji*. Racibórz: Państwowa Wyższa Szkoła Zawodowa, s. 99–108.
18. Kolářová, I. (2011): Univerbizace. In: Štícha, F. (ed.): *Kapitoly z české gramatiky*. Praha: Academia, s. 273–410.
19. Martinčová, O. a kol (2004). *Nová slova v češtině. Slovník neologismů 2*. Praha: Academia
20. Nekvapil, J. (1985): Několik poznámek k využívání formantu -ák ve slanzích a ve spisovném jazyce. In: *NŘ*, roč. 68, 1985, č. 1
21. Prošek, M. (2005): O jednom typu univerbizovaných pojmenování míst na -ák. In: *NŘ*, roč. 88, 2005, č. 2

Jazykové korpusy

22. Benešová, L. – Křen, M. – Waclawičová, M. (2013): *ORAL2013: reprezentativní korpus neformální mluvené češtiny*. Ústav Českého národního korpusu FF UK, Praha. Dostupné z: <http://www.korpus.cz>
23. Křen, M. – Bartoň, T. – Cvrček, V. – Hnátková, M. – Jelínek, T. – Kocek, J. – Novotná, R. – Petkevič, V. – Procházka, P. – Schmiedtová, V. – Skoumalová, H. (2010): *SYN2010: žánrově vyvážený korpus psané češtiny*. Ústav Českého národního korpusu FF UK, Praha. Dostupné z: <http://www.korpus.cz>